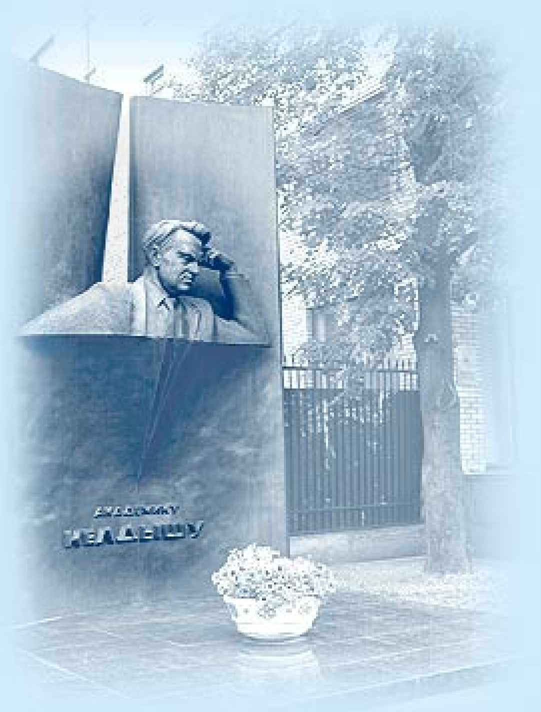

ISSN 2071-2898 (Print)
ISSN 2071-2901 (Online)

S.V. Ershov, M.S. Kopylov,
A.G. Voloboy

Separable Optimal Weights for Bidirectional Ray Tracing with Photon Maps while Mixing 3 Strategies

Recommended form of bibliographic references: Ershov S.V., Kopylov M.S., Voloboy A.G. Separable Optimal Weights for Bidirectional Ray Tracing with Photon Maps while Mixing 3 Strategies // Keldysh Institute Preprints. 2024. No. 82. 40 p. **https://doi.org/10.20948/prepr-2024-82-e**
https://library.keldysh.ru/preprint.asp?id=2024-82&lg=e

О р д е н а Л е н и н а
ИНСТИТУТ ПРИКЛАДНОЙ МАТЕМАТИКИ
имени М.В. Келдыша
Р о с с и й с к о й а к а д е м и и н а у к

S.V. Ershov, M.S. Kopylov, A.G. Voloboy

Separable Optimal Weights for Bidirectional Ray Tracing with Photon Maps while Mixing 3 Strategies

Москва — 2024

С.В. Ершов, М.С. Копылов, А.Г. Волобой

Разделяющиеся оптимальные веса для случая трех стратегий в двунаправленной трассировке лучей с фотонными картами

Шум в итоговом изображении присущ широко используемой двунаправленной стохастической трассировке лучей с фотонными картами. Для снижения шума используется выборка с множественной значимостью, объединяющая результаты различных стратегий с весами. В статье мы исследуем случай разделяющихся весов, выводим и решаем систему интегральных уравнений, определяющих оптимальные веса. Она качественно отличается от ранее исследованного случая весов общего вида и приводит к более простым и численно устойчивым выражениям. Они допускают решение в виде алгебраической формулы, включающей интегралы известных функций, которые могут быть вычислены при трассировке лучей. Исследована роль ее интегральных членов и показано, что они не могут быть заменены простыми эвристиками, а должны быть вычислены с высокой точностью.

Ключевые слова: стохастическая трассировка лучей, фотонные карты, множественная выборка по значимости, оптимальные веса.

S.V. Ershov, M.S. Kopylov, A.G. Voloboy

Separable Optimal Weights for Bidirectional Ray Tracing with Photon Maps while Mixing 3 Strategies

Noise in the resultant image is inherent in the widely used bidirectional stochastic ray tracing with photon maps. The noise can be reduced by the Multiple Importance Sampling which combines results of different strategies with weights. In this paper we investigate the separable weights, derive and solve the system of integral equations that determine the optimal weights. It has several qualitative differences from the previously investigated case of general weights and results in more simple and numerically robust expressions. As before, they can be solved in the form of an algebraic formula that includes several integrals of the known functions that can be calculated in ray tracing. The integral terms entering those expressions are of primary importance and do not admit any simple heuristics but must be calculated accurately.

Key words: stochastic ray tracing, photon maps, multiple importance sampling, optimal weights.

Contents

1 Introduction

Modern realistic computer graphics is based on lighting simulation. Nowadays there are a plenty of methods used for light transport simulation [1] most of which use the Metropolis approach [2] or Monte Carlo ray tracing (MCRT) [3] and their bidirectional modifications. Among them, the bidirectional Monte-Carlo ray tracing with photon maps (BDPM) [4, 5] is effective and widely used method. Stochastic noise is one of the main problems of all methods and many works are devoted to the task of its reduction [6–8]. The majority of them proposed to apply the multiple importance sampling (MIS) technique for decreasing the noise [7–9]. They are based on the theorem [10] assuming independent samples, i.e. full paths connecting the source and the camera. The classic Veach results assume two different formulae to calculate the optimal weights named the "balance heuristic" and the "power heuristic". This is because Veach minimized not the variance itself but rather its bounds. This approach has been improved in [11, 12]. There the resulting formulae for the optimal weights consist of the parts present in the "balance heuristic" with scale coefficients α calculated from the linear system.

But the situation is different for BDPM. Here the same light path is merged with many camera paths and vice versa. Therefore the samples are not independent. In [11] the bi-directional path tracing (BDPT) was considered while we use the bi-directional ray tracing with photon maps (BDPM). Roughly, they trace one light path of the desired number of segments and one camera path with the desired depth. These depths determine the strategy as it is in the classic MIS. The ends of these paths are connected with a straight segment. If it is not occluded, then we increment the accumulated pixel luminance by its "importance function", otherwise the increment is 0. After that the pair of light and camera paths is discarded and a new one is generated. In BDPM, the process is basically different. We trace N_B camera rays and N_F light rays, then each camera path is checked against each light path. If they have close vertices, we get the joint path and increment the accumulated pixel luminance by its "importance function". As a result, the BDPT operates independent samples while BDPM operates dependent samples.

The problem of the optimal weights in a limited MIS for mixing contributions from two first camera vertices was considered in [13, 14]. There the main idea of calculation of the optimal weights for BDPM was stated: while the famous heuristic for the optimal weights in MCRT [10] obey a system of algebraic equations, in BDPM they obey the system of integral equations. Sometimes mixing two strategies can be not enough. It is possible that only the third (or later) vertex has smooth BDF (Bidirectional scattering Distribution Function which describes optical properties of a surface or an object in scene) where gathering of illumination will be good. So mixing of three or more strategies is an interesting case.

In [15] we demonstrated that the case of three or more strategies is qualitatively different from mixing two strategies. The main reason is that while for mixing two strategies we had only one variable weight, now we have two different families of weights, each with its own normalization conditions. The first family consists of two weights, one of them is dependent. The second family consists of three weights, any two of them can be considered as independent and the remaining one is dependent on them. In case of mixing more strategies there will be more families of weights. All independent weights from all families are coupled in the common system of integral equations while for two strategies we had one weight and one integral equation. The resulting system of integral equations admits local approximation and reduces to the system of algebraic equations because the integral terms can be neglected. The formula used for two strategies allows for an "intuitive induction" to cases of three and more strategies. So one can imagine what it will look like.

In [15] we had calculated the optimal weights for MIS in BDPM, when mixed are three strategies (merge the camera and light rays at the first, or at the second, or at the third diffuse event of camera ray). It was a "limited MIS" because the number of strategies is limited, and the weights are assumed to depend only on the three first vertices of the joint ray path.

The weights satisfy a system of integral equations that admit solution in a closed form, i.e. the weight are algebraic expressions which include some definite integrals. There are, however, two problems. First, the expressions used are very sophisticated and do not admit any intuitive understanding which would allow to analyze the role of the scene parameters etc. That is they admit only numerical calculation. And here is the second difficulty. The weights are a solution of a 2x2 linear algebraic system, whose matrix and right hand side are integrals whose integrands are calculated via solution of another 2x2 linear algebraic system, whose matrix and right hand side are integrals. For a general complex scene those integrals only admit a Monte-Carlo calculation. Meanwhile, for finite calculation time it may give noisy, thus not very accurate, results. In case the matrices have a small determinant, this inaccuracy may result in a big error in the solution. Therefore one must calculate all the integrals in the weight expressions with high accuracy. For Monte-Carlo this usually means long calculation time. As a result, calculation of the weights is expensive and this may cancel all the gain which using these weights in the BDPM may give.

2 Separable weights

2.1 The idea of the separable weights

Calculation of weights in [15] is very sophisticated and includes solution of a linear system whose matrix is to be calculated numerically, and, generally, with MCRT. Therefore we know it with a rather limited accuracy. Since its stability (against perturbations of the matrix etc.) had not been investigated it is possible the solution will be sensitive to the numeric errors and we shall obtain the optimal weights with substantial inaccuracy. Or, we should calculate the integrals with very high accuracy. This may render the calculations so expensive that this kills all the gain from weighting strategies. At last, the formulae are not intuitive and it is difficult to comprehend how they work and how various characteristics of the scene affect the weights.

In this work we investigate a “stratification” of the problem based on the intuitive idea. Namely, let first suppose that for each ray only single strategy is in effect, that is, we either join the camera and the light paths at the 1st, or at the 2nd, or at the 3rd vertex of the camera ray. The weight is then 0 for all strategies but one, for which it is 1. The “stratification” splits the problem of choosing a vertex to merge into two sequential choices. First, we decide whether to merge at the 1st vertex or anywhere beyond (for the case of 3 strategies this means at the 2nd or at the 3rd). In case of the last choice, the second step is to decide, whether we merge rays at the 2nd or at the 3rd vertex. This scheme is intuitively simple, understandable and easily extends to n strategies.

If we allow for the fractional weights, the situation is a bit more sophisticated, but still we can calculate the weights “separately”, first the fraction among the 0th strategy and all the rest and then between the 1st strategy and all the rest and so on. As a result, at each step we operate relative weights among merging at the given vertex and mixing anywhere beyond it.

This problem leads to simpler and more intuitive equations for the weights. Usually, they are stable against perturbations of calculating the coefficients of the equations. Also, these equations admit "human understanding" of which scene parameters affect the weights and how.

2.2 Formal statement of the problem

Here and below we base upon [15], retain the notations and calculation of the BDPM noise in case of mixing strategies and only use separable weights.

Intuitively, the contribution of the forward MCRT ray at the merging vertex must not seriously depend on its far prehistory (where it came from). So it seems natural to believe that one can take $w_{3,0}(\boldsymbol{z}_0, \boldsymbol{z}_1, \boldsymbol{z}_2) = w_{3,0}(\boldsymbol{z}_0, \boldsymbol{z}_1)$. But because of the normalization condition $w_{3,0} + w_{3,1} + w_{3,2} = 1$ the sum $w_{3,1} + w_{3,2} \equiv w_{3,1+2}$ must also be independent of $\boldsymbol{z}_2$!

Obviously, simple dropping this last vertex from both $w_{3,1}$ and $w_{3,2}$ will not work, because then the choice whether to merge the rays at $\boldsymbol{z}_2$ would be independent from that $\boldsymbol{z}_2$. We therefore require that these weights do depend on three points while their sum is independent of $\boldsymbol{z}_2$, that is

$$\begin{aligned} w_{3,0}(\boldsymbol{z}_0, \boldsymbol{z}_1, \boldsymbol{z}_2) &= w_{3,0}(\boldsymbol{z}_0, \boldsymbol{z}_1) \\ w_{3,1}(\boldsymbol{z}_0, \boldsymbol{z}_1, \boldsymbol{z}_2) &= w_{3,1+2}(\boldsymbol{z}_0, \boldsymbol{z}_1)\omega_{3,1}(\boldsymbol{z}_0, \boldsymbol{z}_1, \boldsymbol{z}_2) \\ w_{3,2}(\boldsymbol{z}_0, \boldsymbol{z}_1, \boldsymbol{z}_2) &= w_{3,1+2}(\boldsymbol{z}_0, \boldsymbol{z}_1)\omega_{3,2}(\boldsymbol{z}_0, \boldsymbol{z}_1, \boldsymbol{z}_2) \end{aligned} \tag{1}$$

with separate normalization

$$\begin{aligned} w_{3,0}(\boldsymbol{z}_0, \boldsymbol{z}_1) + w_{3,1+2}(\boldsymbol{z}_0, \boldsymbol{z}_1) &= 1 \\ \omega_{3,1}(\boldsymbol{z}_0, \boldsymbol{z}_1, \boldsymbol{z}_2) + \omega_{3,2}(\boldsymbol{z}_0, \boldsymbol{z}_1, \boldsymbol{z}_2) &= 1 \end{aligned}$$

In view of our idea to neglect "far history", it is natural to use the same weight for the forward MCRT rays of 2 and of 3 segments when hitting $\boldsymbol{z}_0$. In [15], the former had weight $w_{2,0}$ while the latter had the weight $w_{3,0}$. Now, therefore, they are identified: $w_{2,0} = w_{3,0}$. It was impossible for the weights used in [15] because there the function $w_{3,0}$ has on three arguments while $w_{2,0}$ has two. But for our separable weights, when $w_{3,0} = w_{3,0}(\boldsymbol{z}_0, \boldsymbol{z}_1)$, this is possible.

Since $w_{2,1} = 1 - w_{2,0}$ and $w_{3,1+2} = 1 - w_{3,0}$, the rest weights are also equal: $w_{2,1} = w_{3,1+2}$. Thus we now have a single family of weights instead of two families in [15]. So, in current work we start with the general formulae from [15] and substitute there (1) and

$$
\begin{aligned}
w_{2,0} &= w_{3,0} \\
w_{2,1} &= w_{3,1+2}
\end{aligned} \tag{2}
$$

Formally, as explained in [15], we can use arbitrary weights provided they satisfy normalization conditions (the last equation of Section 3 in [15]). Regardless of the reasons to use weights of the form (1)–(2), they do satisfy those conditions and thus we can use weights of such type. Since this is a constraint reducing the number of degrees of freedom, the new weights are "less optimal" than those from [15]. On the other hand, they are simpler and more stable.

3 BDPM noise for separable weights

Here and below all calculations are for one pixel like in [15]. For the sake of simplicity, the total flux of all scene lights is assumed 1 to not bother about scaling between the density of photons and irradiance. We also assume that the light source is a point one and that it is unique in the scene.

The material from Sections 3 and 4 of [15] applies here without any changes so we do not reproduce it here.

In BDPM the variance V of the pixel luminance calculated from N_F forward rays and N_B backward rays started from the same pixel obeys the general law [6]:

$$
\begin{aligned}
V &= \frac{1}{N_F N_B}\left(\langle\langle C^2\rangle\rangle - \langle\langle C\rangle\rangle^2\right) + \frac{1-N_B^{-1}}{N_F}\left(\langle\langle C\rangle_B^2\rangle_F - \langle\langle C\rangle\rangle^2\right) \\
&\quad + \frac{1-N_F^{-1}}{N_B}\left(\langle\langle C\rangle_F^2\rangle_B - \langle\langle C\rangle\rangle^2\right)
\end{aligned} \tag{3}
$$

Here C is the contribution of the merged forward and backward rays to the pixel luminance (if they do not intersect, $C = 0$) given by eq. (4) of [15], $\langle\cdot\rangle_B$ is the averaging over the backward rays ensemble for the fixed light ray and $\langle\cdot\rangle_F$ is the averaging over the forward rays ensemble for the fixed camera ray. Notice the linear term $\langle\langle C\rangle\rangle$ is independent from the order of averaging so we drop subscripts here. It is also independent from weights, while $\langle\langle C^2\rangle\rangle$ and $\langle\langle C\rangle_F^2\rangle_B$ depend on them.

Averaging over the ensemble of light paths resp. camera paths is

$$
\langle\cdot\rangle_F = \int (\cdot) p_F(\boldsymbol{x}_1, ..., \boldsymbol{x}_n, ...) d\boldsymbol{x}_1 \cdots d\boldsymbol{x}_n \cdots \tag{4}
$$

$$
\langle\cdot\rangle_B = \int (\cdot) p_B(\boldsymbol{y}_0, \boldsymbol{y}_1, \boldsymbol{y}_2) d\boldsymbol{y}_1 d\boldsymbol{y}_2 \tag{5}
$$

where $(\boldsymbol{y}_{-1}, \boldsymbol{y}_0, \boldsymbol{y}_1, ...)$ is the camera path and $(\boldsymbol{x}_0, \boldsymbol{x}_1, \boldsymbol{x}_2, ...)$ is the light path, p_F and p_B are the probability densities of the light and camera paths (the latter for the fixed given pixel). Since we assume that the forward MCRT (FMCRT) uses Russian roulette to kill rays while keeps ray energy, p_F is not normalized (i.e. its integral is not 1). The fixed points $\boldsymbol{x}_0$ and $\boldsymbol{y}_{-1}$ are not included in the averaging. The point $\boldsymbol{y}_0$ depends on pixel, but for given pixel it is also fixed and thus is not also included in the averaging. The point $\boldsymbol{y}_{-1}$ (camera origin) is always the same for all pixels, so we frequently drop it from the list of arguments of a function.

The formulae for various averages of the path contribution from [15] apply to arbitrary weights $w_{3,m}$ and $w_{2,m}$ and therefore are valid for their particular form (1)–(2).

Hereafter $\boldsymbol{z}_j$ are the vertices of the joint path numbered from the camera end. It can be taken from either the camera or the light subpath depending on the intersection type. We drop camera origin $\boldsymbol{z}_{-1} = \boldsymbol{y}_{-1}$ from this set because it is fixed point. The concept of $\boldsymbol{z}_j$ is also described in detail in [15].

3.1 Cross term and forward MCRT term

Substituting these separable weights into the equations (10) and (12) from [15] we can write the double averages $\langle\langle C^2\rangle\rangle$ and $\langle\langle C\rangle_B^2\rangle_F$ entering the noise formula as

$$\begin{aligned}\langle\langle C^2\rangle\rangle \;&\approx\; \frac{1}{S}\int w_{3,0}^2(\boldsymbol{z}_0,\boldsymbol{z}_1) f^2(\overrightarrow{\boldsymbol{z}_1\boldsymbol{z}_0}, \overrightarrow{\boldsymbol{y}_{-1}\boldsymbol{z}_0}, \boldsymbol{z}_0) L(\boldsymbol{z}_0,\boldsymbol{z}_1)\left|(\boldsymbol{n}(\boldsymbol{z}_0)\cdot\overrightarrow{\boldsymbol{z}_1\boldsymbol{z}_0})\right| s(\boldsymbol{z}_0,\boldsymbol{z}_1) d\boldsymbol{z}_1 \\ &+\frac{1}{S}\int w_{3,1}^2(\boldsymbol{z}_0,\boldsymbol{z}_1,\boldsymbol{z}_2) E(\boldsymbol{z}_0) f(\overrightarrow{\boldsymbol{z}_2\boldsymbol{z}_1}, \overrightarrow{\boldsymbol{z}_0\boldsymbol{z}_1}, \boldsymbol{z}_1)\rho(\boldsymbol{z}_0,\boldsymbol{z}_1,\boldsymbol{z}_2) d\boldsymbol{z}_1 d\boldsymbol{z}_2 \\ &+\frac{1}{S}\int w_{3,2}^2(\boldsymbol{z}_0,\boldsymbol{z}_1,\boldsymbol{z}_2) E(\boldsymbol{z}_0,\boldsymbol{z}_1)\frac{b(\boldsymbol{z}_1,\boldsymbol{z}_2)}{L(\boldsymbol{z}_1,\boldsymbol{z}_2)}\rho(\boldsymbol{z}_0,\boldsymbol{z}_1,\boldsymbol{z}_2) d\boldsymbol{z}_1 d\boldsymbol{z}_2\end{aligned}$$

and

$$\langle\langle C\rangle_B^2\rangle_F \approx \frac{1}{S}\int w_{3,0}^2(\boldsymbol{z}_0,\boldsymbol{z}_1) f^2(\overrightarrow{\boldsymbol{z}_1\boldsymbol{z}_0}, \overrightarrow{\boldsymbol{y}_{-1}\boldsymbol{z}_0}, \boldsymbol{z}_0) L(\boldsymbol{z}_0,\boldsymbol{z}_1)\left|(\boldsymbol{n}(\boldsymbol{z}_0)\cdot\overrightarrow{\boldsymbol{z}_0\boldsymbol{z}_1})\right| s(\boldsymbol{z}_0,\boldsymbol{z}_1) d\boldsymbol{z}_1$$

where the "≈" means "up to $O(1)$" (notice we assume the area of the "integration sphere" S is small!),

$$s(\overrightarrow{\boldsymbol{x}\boldsymbol{y}}) \;\equiv\; \frac{\left|(\overrightarrow{\boldsymbol{x}\boldsymbol{y}}\cdot\boldsymbol{n}(\boldsymbol{y}))\right|}{\left|\boldsymbol{x}-\boldsymbol{y}\right|^2}$$

where $\boldsymbol{n}(\boldsymbol{y})$ is the local normal at the point $\boldsymbol{y}$ and $\overrightarrow{\boldsymbol{xy}}$ is the unit vector from $\boldsymbol{x}$ to $\boldsymbol{y}$ relates the differentials of solid angle and of surface area

$$d^2(\overrightarrow{\boldsymbol{xy}}) \;=\; s(\boldsymbol{x},\boldsymbol{y})d\boldsymbol{y} \tag{6}$$

Then, $f(\boldsymbol{v},\boldsymbol{u},\boldsymbol{x})$ is BDF (in luminance units) of the surface point $\boldsymbol{x}$ for illumination direction $\boldsymbol{v}$ and viewing direction $\boldsymbol{u}$, and $E(\boldsymbol{y}_0,...,\boldsymbol{y}_{i-1})$ is the energy of the camera ray before hitting $\boldsymbol{y}_i$. This energy (or transmission factor in [5] terms) is defined as usual: it is 1 just after leaving the camera, i.e. $E(\boldsymbol{y}_{-1}) = 1$ and then

$$\begin{aligned} E(\boldsymbol{x}_0) &= \mu(\overrightarrow{\boldsymbol{x}_{-1}\boldsymbol{x}_0},\boldsymbol{x}_0), \\ E(\boldsymbol{x}_0,...,\boldsymbol{x}_m) &= \mu(\overrightarrow{\boldsymbol{x}_{m-1}\boldsymbol{x}_m},\boldsymbol{x}_m)E(\boldsymbol{x}_0,...,\boldsymbol{x}_{m-1}), \qquad m = 1,2,... \end{aligned} \tag{7}$$

where

$$\mu(\boldsymbol{u},\boldsymbol{x}) \equiv \int f(\boldsymbol{v},\boldsymbol{u},\boldsymbol{x})\left|(\boldsymbol{v}\cdot\boldsymbol{n}(\boldsymbol{x}))\right| d^2\boldsymbol{v}, \tag{8}$$

$\boldsymbol{n}(\boldsymbol{x})$ being the local normal at the point $\boldsymbol{x}$. Using formula (40) from [15] we rewrite the ρ in the following form:

$$\begin{aligned} \rho(\boldsymbol{z}_0,\boldsymbol{z}_1,\boldsymbol{z}_2) &\equiv E(\boldsymbol{z}_0)L(\boldsymbol{z}_1,\boldsymbol{z}_2) \\ &\quad \times f(\overrightarrow{\boldsymbol{z}_2\boldsymbol{z}_1},\overrightarrow{\boldsymbol{z}_0\boldsymbol{z}_1},\boldsymbol{z}_1)(\boldsymbol{n}(\boldsymbol{z}_1)\cdot\overrightarrow{\boldsymbol{z}_2\boldsymbol{z}_1})s(\boldsymbol{z}_1,\boldsymbol{z}_2)p_B(\boldsymbol{z}_0,\boldsymbol{z}_1) \\ &= E(\boldsymbol{z}_0,\boldsymbol{z}_1)L(\boldsymbol{z}_1,\boldsymbol{z}_2) \\ &\quad \times \tilde{f}(\overrightarrow{\boldsymbol{z}_2\boldsymbol{z}_1},\overrightarrow{\boldsymbol{z}_0\boldsymbol{z}_1},\boldsymbol{z}_1)(\boldsymbol{n}(\boldsymbol{z}_1)\cdot\overrightarrow{\boldsymbol{z}_2\boldsymbol{z}_1})s(\boldsymbol{z}_1,\boldsymbol{z}_2)p_B(\boldsymbol{z}_0,\boldsymbol{z}_1) \end{aligned} \tag{9}$$

where $\tilde{f}$ is the normalized BDF. The integral

$$b(\boldsymbol{x},\boldsymbol{y}) \equiv \int f^2(\overrightarrow{\boldsymbol{zy}},\overrightarrow{\boldsymbol{yx}},\boldsymbol{y})L(\boldsymbol{y},\boldsymbol{z})\left|(\boldsymbol{n}(\boldsymbol{y})\cdot\overrightarrow{\boldsymbol{zy}})\right| s(\boldsymbol{y},\boldsymbol{z})d\boldsymbol{z} \tag{10}$$

is independent of weights and is much similar to the integral which gives diffuse luminance of $\boldsymbol{y}$ towards $\boldsymbol{x}$

$$L_d(\boldsymbol{x},\boldsymbol{y}) \equiv \int f(\overrightarrow{\boldsymbol{zy}},\overrightarrow{\boldsymbol{yx}},\boldsymbol{y})L(\boldsymbol{y},\boldsymbol{z})\left|(\boldsymbol{n}(\boldsymbol{y})\cdot\overrightarrow{\boldsymbol{zy}})\right| s(\boldsymbol{y},\boldsymbol{z})d\boldsymbol{z} \tag{11}$$

All notations and terms are the same as in [15] where they had been explained in detail.

3.2 Backward MCRT term

The third term in the noise equation, the Backward MCRT (BMCRT) term, can be retained in the form of eq. (15) from [15] but it is inconvenient. Instead, we first rewrite $\langle C\rangle_F$ in a slightly different form so that

$$\begin{aligned}\langle C\rangle_F - \langle\langle C\rangle\rangle &= w_{2,1}(\boldsymbol{z}_0,\boldsymbol{z}_1)E(\boldsymbol{z}_0)L_0(\boldsymbol{z}_0,\boldsymbol{z}_1) - \hat{F}_1 + E(\boldsymbol{z}_0)G_1(\boldsymbol{z}_1)\\ &\quad - \int \left(G_1(\boldsymbol{z}_1)+G_2(\boldsymbol{z}_1)\right) f(\overrightarrow{\boldsymbol{z}_0\boldsymbol{z}_1},\overrightarrow{\boldsymbol{y_{-1}}\boldsymbol{z}_0},\boldsymbol{z}_0)\\ &\quad\quad \times \left|(\boldsymbol{n}(\boldsymbol{z}_0)\cdot\overrightarrow{\boldsymbol{z}_1\boldsymbol{z}_0})\right| s(\boldsymbol{z}_0,\boldsymbol{z}_1)d\boldsymbol{z}_1\\ &\quad + w_{3,2}(\boldsymbol{z}_0,\boldsymbol{z}_1,\boldsymbol{z}_2)E(\boldsymbol{z}_0,\boldsymbol{z}_1)L(\boldsymbol{z}_1,\boldsymbol{z}_2)\end{aligned}$$

Here $L_0(\boldsymbol{z}_0,\boldsymbol{z}_1)$ is the luminance due to direct illumination, and (like in (13)–(14) of [15])

$$\begin{aligned}\hat{F}_m &\equiv \int w_{2,m}(\boldsymbol{z}_0,\boldsymbol{z}_1)L_0(\boldsymbol{z}_0,\boldsymbol{z}_1)f(\overrightarrow{\boldsymbol{z}_0\boldsymbol{z}_1},\overrightarrow{\boldsymbol{y_{-1}}\boldsymbol{z}_0},\boldsymbol{z}_0)\\ &\quad\times \left|(\boldsymbol{n}(\boldsymbol{z}_0)\cdot\overrightarrow{\boldsymbol{z}_1\boldsymbol{z}_0})\right| s(\boldsymbol{z}_0,\boldsymbol{z}_1)d\boldsymbol{z}_1 \qquad (12)\\ G_m(\boldsymbol{z}_1) &\equiv \int w_{3,m}(\boldsymbol{z}_0,\boldsymbol{z}_1,\boldsymbol{z}_2)L(\boldsymbol{z}_1,\boldsymbol{z}_2)f(\overrightarrow{\boldsymbol{z}_1\boldsymbol{z}_2},\overrightarrow{\boldsymbol{z}_0\boldsymbol{z}_1},\boldsymbol{z}_1)\\ &\quad\times \left|(\boldsymbol{n}(\boldsymbol{z}_1)\cdot\overrightarrow{\boldsymbol{z}_2\boldsymbol{z}_1})\right| s(\boldsymbol{z}_1,\boldsymbol{z}_2)d\boldsymbol{z}_2 \qquad (13)\end{aligned}$$

For the separable weights (1)–(2), the above functions become

$$\hat{F}_1 \equiv \int w_{3,1+2}(\boldsymbol{z}_0,\boldsymbol{z}_1)L_0(\boldsymbol{z}_0,\boldsymbol{z}_1)f(\overrightarrow{\boldsymbol{z}_0\boldsymbol{z}_1},\overrightarrow{\boldsymbol{y_{-1}}\boldsymbol{z}_0},\boldsymbol{z}_0)\left|(\boldsymbol{n}(\boldsymbol{z}_0)\cdot\overrightarrow{\boldsymbol{z}_1\boldsymbol{z}_0})\right| s(\boldsymbol{z}_0,\boldsymbol{z}_1)d\boldsymbol{z}_1$$

and

$$G_m(\boldsymbol{z}_1) = w_{3,1+2}(\boldsymbol{z}_0,\boldsymbol{z}_1)\Gamma_m(\boldsymbol{z}_0,\boldsymbol{z}_1)$$

where

$$\begin{aligned}\Gamma_m(\boldsymbol{z}_0,\boldsymbol{z}_1) &\equiv \int \omega_{3,m}(\boldsymbol{z}_0,\boldsymbol{z}_1,\boldsymbol{z}_2)L(\boldsymbol{z}_1,\boldsymbol{z}_2)\\ &\quad\times f(\overrightarrow{\boldsymbol{z}_1\boldsymbol{z}_2},\overrightarrow{\boldsymbol{z}_0\boldsymbol{z}_1},\boldsymbol{z}_1)\left|(\boldsymbol{n}(\boldsymbol{z}_1)\cdot\overrightarrow{\boldsymbol{z}_2\boldsymbol{z}_1})\right| s(\boldsymbol{z}_1,\boldsymbol{z}_2)d\boldsymbol{z}_2\\ \widetilde{\Gamma}_m(\boldsymbol{z}_0,\boldsymbol{z}_1) &\equiv \int \omega_{3,m}(\boldsymbol{z}_0,\boldsymbol{z}_1,\boldsymbol{z}_2)L(\boldsymbol{z}_1,\boldsymbol{z}_2) \qquad (14)\\ &\quad\times \tilde{f}(\overrightarrow{\boldsymbol{z}_1\boldsymbol{z}_2},\overrightarrow{\boldsymbol{z}_0\boldsymbol{z}_1},\boldsymbol{z}_1)\left|(\boldsymbol{n}(\boldsymbol{z}_1)\cdot\overrightarrow{\boldsymbol{z}_2\boldsymbol{z}_1})\right| s(\boldsymbol{z}_1,\boldsymbol{z}_2)d\boldsymbol{z}_2\\ &= \frac{E(\boldsymbol{z}_0)}{E(\boldsymbol{z}_0,\boldsymbol{z}_1)}\Gamma_m(\boldsymbol{z}_0,\boldsymbol{z}_1)\end{aligned}$$

and then

$$\begin{aligned}\langle C\rangle_F - \langle\langle C\rangle\rangle &= w_{3,1+2}(\boldsymbol{z}_0,\boldsymbol{z}_1)E(\boldsymbol{z}_0)L_0(\boldsymbol{z}_0,\boldsymbol{z}_1) - \hat{F}_1 \\ &\quad + E(\boldsymbol{z}_0)w_{3,1+2}(\boldsymbol{z}_0,\boldsymbol{z}_1)\Gamma_1(\boldsymbol{z}_0,\boldsymbol{z}_1) \\ &\quad - \int w_{3,1+2}(\boldsymbol{z}_0,\boldsymbol{z}_1)\left(\Gamma_1(\boldsymbol{z}_0,\boldsymbol{z}_1)+\Gamma_2(\boldsymbol{z}_0,\boldsymbol{z}_1)\right) \\ &\qquad \times f(\overrightarrow{\boldsymbol{z}_0\boldsymbol{z}_1}, \overrightarrow{\boldsymbol{y_{-1}}\boldsymbol{z}_0}, \boldsymbol{z}_0)\left|(\boldsymbol{n}(\boldsymbol{z}_0)\cdot\overrightarrow{\boldsymbol{z}_1\boldsymbol{z}_0})\right| s(\boldsymbol{z}_0,\boldsymbol{z}_1)d\boldsymbol{z}_1 \\ &\quad + w_{3,1+2}(\boldsymbol{z}_0,\boldsymbol{z}_1)\omega_{3,2}(\boldsymbol{z}_0,\boldsymbol{z}_1,\boldsymbol{z}_2)E(\boldsymbol{z}_0,\boldsymbol{z}_1)L(\boldsymbol{z}_1,\boldsymbol{z}_2)\end{aligned}$$

Obviously,

$$\Gamma_1(\boldsymbol{z}_0,\boldsymbol{z}_1) + \Gamma_2(\boldsymbol{z}_0,\boldsymbol{z}_1) = L_d(\boldsymbol{z}_0,\boldsymbol{z}_1)$$

and introducing

$$\begin{aligned}\widetilde{\Phi}(\boldsymbol{z}_0) \equiv \int & w_{3,1+2}(\boldsymbol{z}_0,\boldsymbol{z}_1)L(\boldsymbol{z}_0,\boldsymbol{z}_1)\tilde{f}(\overrightarrow{\boldsymbol{z}_0\boldsymbol{z}_1}, \overrightarrow{\boldsymbol{y_{-1}}\boldsymbol{z}_0}, \boldsymbol{z}_0) \\ & \times\left|(\boldsymbol{n}(\boldsymbol{z}_0)\cdot\overrightarrow{\boldsymbol{z}_1\boldsymbol{z}_0})\right| s(\boldsymbol{z}_0,\boldsymbol{z}_1)d\boldsymbol{z}_1\end{aligned} \tag{15}$$

(where the full luminance $L(\boldsymbol{z}_0,\boldsymbol{z}_1) = L_0(\boldsymbol{z}_0,\boldsymbol{z}_1) + L_d(\boldsymbol{z}_0,\boldsymbol{z}_1)$) we arrive at

$$\begin{aligned}\langle C\rangle_F - \langle\langle C\rangle\rangle &= E(\boldsymbol{z}_0)\left(L(\boldsymbol{z}_0,\boldsymbol{z}_1)w_{3,1+2}(\boldsymbol{z}_0,\boldsymbol{z}_1) - \widetilde{\Phi}\right) \\ &\quad + E(\boldsymbol{z}_0,\boldsymbol{z}_1)\left(L(\boldsymbol{z}_1,\boldsymbol{z}_2)\omega_{3,2}(\boldsymbol{z}_0,\boldsymbol{z}_1,\boldsymbol{z}_2) - \widetilde{\Gamma}_2(\boldsymbol{z}_0,\boldsymbol{z}_1)\right) w_{3,1+2}(\boldsymbol{z}_0,\boldsymbol{z}_1)\end{aligned}$$

The expression $\langle\langle C\rangle_F^2\rangle_B - \langle\langle C\rangle\rangle^2 = \langle(\langle C\rangle_F - \langle\langle C\rangle\rangle)^2\rangle_B$ in the last, BMCRT term of the noise equation is by definition the average of the square of the above difference:

$$\langle\langle C\rangle_B^2\rangle_F - \langle\langle C\rangle\rangle^2 = \int\left(\int\left(\langle C\rangle_F - \langle\langle C\rangle\rangle\right)^2 p_B(\boldsymbol{z}_0,\boldsymbol{z}_1,\boldsymbol{z}_2)d\boldsymbol{z}_2\right)d\boldsymbol{z}_0 d\boldsymbol{z}_1$$

The integrals $\int\cdots p_B(\boldsymbol{z}_1,\boldsymbol{z}_2)d\boldsymbol{z}_2$ are in fact the averaging over BMCRT rays expressed by the integral over all rays of $\overrightarrow{\boldsymbol{z}_1\boldsymbol{z}_2}$ directions, including rays missed scene surfaces. For the "spatial" integral $\int\cdots d\boldsymbol{z}_2$ this can be treated in the following way (Fig. 1): it runs over the large black (totally absorbing) bounding sphere (enclosing the scene). If the ray from $\boldsymbol{z}_1$ point intersects a scene surface, then $\boldsymbol{z}_2$ in the integrand is this (first) ray intersection. Otherwise $\boldsymbol{z}_2$ is the point on the black bounding sphere. On the black sphere all luminance components ($L(\boldsymbol{z}_1,\boldsymbol{z}_2)$, $L_0(\boldsymbol{z}_1,\boldsymbol{z}_2)$ and $L_d(\boldsymbol{z}_1,\boldsymbol{z}_2)$) vanish, as well as BDF. The term $b(\boldsymbol{z}_1,\boldsymbol{z}_2)$

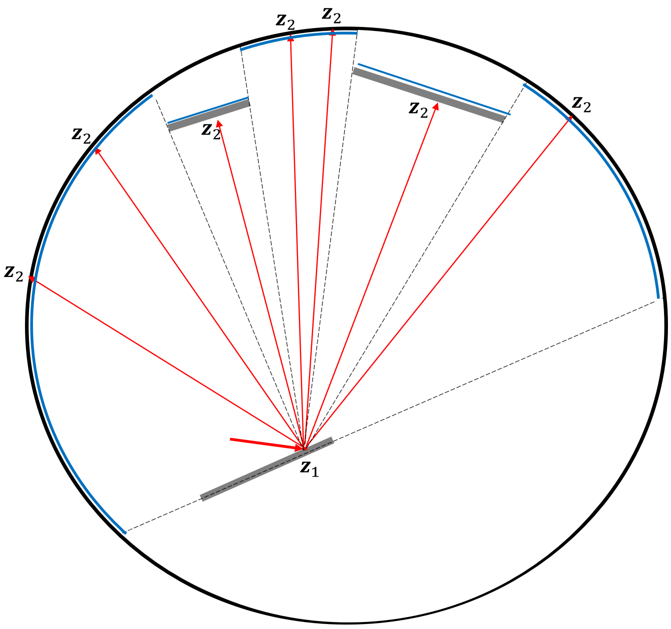

Figure 1: 2D cross-section of a scene. The integration domain over $\boldsymbol{z}_2$ for the given $\boldsymbol{z}_1$ is thin blue line consisting of five parts . Thick gray segments are the cross-sections of the scene surfaces. Black circle is the cross-section of the black bounding sphere. The arrows are the camera rays. Dashed lines are guides.

vanishes too, because it includes BDF at $\boldsymbol{z}_2$. The normal is $\boldsymbol{n}(\boldsymbol{z}_2)$ is the internal normal to the bounding sphere. The terms like $\tilde{f}(\overrightarrow{\boldsymbol{z}_2\boldsymbol{z}_1}, \overrightarrow{\boldsymbol{z}_0\boldsymbol{z}_1}, \boldsymbol{z}_1)$ which depend only on the direction $\overrightarrow{\boldsymbol{z}_2\boldsymbol{z}_1}$ and not on the hit point $\boldsymbol{z}_2$ itself, are not 0, as well as the Jacobian $s(\boldsymbol{z}_1, \boldsymbol{z}_2)$ of the space-to-angle transformation. The situation for the integration over $\boldsymbol{z}_1$ for the given $\boldsymbol{z}_0$ is the same.

Therefore, for any g independent of $\boldsymbol{z}_2$ naturally

$$\int g p_B(\boldsymbol{z}_0, \boldsymbol{z}_1, \boldsymbol{z}_2) d\boldsymbol{z}_2 = g \int p_B(\boldsymbol{z}_0, \boldsymbol{z}_1, \boldsymbol{z}_2) d\boldsymbol{z}_2 = g p_B(\boldsymbol{z}_0, \boldsymbol{z}_1)$$

Let us separate in $\langle C \rangle_F - \langle\langle C \rangle\rangle$ the terms which do not depend on $\boldsymbol{z}_2$:

$$\begin{aligned}
\langle C \rangle_F - \langle\langle C \rangle\rangle &= Q_1(\boldsymbol{z}_0, \boldsymbol{z}_1) + Q_2(\boldsymbol{z}_0, \boldsymbol{z}_1) \left(L(\boldsymbol{z}_1, \boldsymbol{z}_2) \omega_{3,2}(\boldsymbol{z}_0, \boldsymbol{z}_1, \boldsymbol{z}_2) - \widetilde{\Gamma}_2(\boldsymbol{z}_0, \boldsymbol{z}_1) \right) \\
Q_1(\boldsymbol{z}_0, \boldsymbol{z}_1) &\equiv E(\boldsymbol{z}_0) \left(L(\boldsymbol{z}_0, \boldsymbol{z}_1) w_{3,1+2}(\boldsymbol{z}_0, \boldsymbol{z}_1) - \widetilde{\Phi} \right) \\
Q_2(\boldsymbol{z}_0, \boldsymbol{z}_1) &\equiv E(\boldsymbol{z}_0, \boldsymbol{z}_1) w_{3,1+2}(\boldsymbol{z}_0, \boldsymbol{z}_1)
\end{aligned}$$

Then, obviously,

$$\int \left(\langle C\rangle_F - \langle\langle C\rangle\rangle\right)^2 p_B(\boldsymbol{z}_0,\boldsymbol{z}_1,\boldsymbol{z}_2)d\boldsymbol{z}_2$$
$$= Q_1^2(\boldsymbol{z}_0,\boldsymbol{z}_1)p_B(\boldsymbol{z}_0,\boldsymbol{z}_1)$$
$$+ Q_2^2(\boldsymbol{z}_0,\boldsymbol{z}_1)\int \Big(L(\boldsymbol{z}_1,\boldsymbol{z}_2)\omega_{3,2}(\boldsymbol{z}_0,\boldsymbol{z}_1,\boldsymbol{z}_2) - \widetilde{\Gamma}_2(\boldsymbol{z}_0,\boldsymbol{z}_1)\Big)^2 p_B(\boldsymbol{z}_0,\boldsymbol{z}_1,\boldsymbol{z}_2)d\boldsymbol{z}_2$$
$$+ 2Q_1(\boldsymbol{z}_0,\boldsymbol{z}_1)Q_2(\boldsymbol{z}_0,\boldsymbol{z}_1)$$
$$\times\left(\int L(\boldsymbol{z}_1,\boldsymbol{z}_2)\omega_{3,2}(\boldsymbol{z}_0,\boldsymbol{z}_1,\boldsymbol{z}_2)p_B(\boldsymbol{z}_0,\boldsymbol{z}_1,\boldsymbol{z}_2)d\boldsymbol{z}_2 - \widetilde{\Gamma}_2(\boldsymbol{z}_0,\boldsymbol{z}_1)p_B(\boldsymbol{z}_0,\boldsymbol{z}_1)\right)$$

Since (see [15], eq. (40))

$$p_B(\boldsymbol{z}_0,\boldsymbol{z}_1,\boldsymbol{z}_2) = p_B(\boldsymbol{z}_0,\boldsymbol{z}_1)\times \tilde{f}(\overrightarrow{\boldsymbol{z}_2\boldsymbol{z}_1},\overrightarrow{\boldsymbol{z}_0\boldsymbol{z}_1},\boldsymbol{z}_1)\left|(\overrightarrow{\boldsymbol{z}_2\boldsymbol{z}_1}\cdot\boldsymbol{n}(\boldsymbol{z}_1))\right| s(\boldsymbol{z}_1,\boldsymbol{z}_2)$$

the last integral is

$$\int L(\boldsymbol{z}_1,\boldsymbol{z}_2)\omega_{3,2}(\boldsymbol{z}_0,\boldsymbol{z}_1,\boldsymbol{z}_2)p_B(\boldsymbol{z}_0,\boldsymbol{z}_1,\boldsymbol{z}_2)d\boldsymbol{z}_2 \;=\; p_B(\boldsymbol{z}_0,\boldsymbol{z}_1)\widetilde{\Gamma}_2(\boldsymbol{z}_0,\boldsymbol{z}_1)$$

Therefore,

$$\int \left(\langle C\rangle_F - \langle\langle C\rangle\rangle\right)^2 p_B(\boldsymbol{z}_0,\boldsymbol{z}_1,\boldsymbol{z}_2)d\boldsymbol{z}_2$$
$$= p_B(\boldsymbol{z}_0,\boldsymbol{z}_1)Q_1^2(\boldsymbol{z}_0,\boldsymbol{z}_1) + Q_2^2(\boldsymbol{z}_0,\boldsymbol{z}_1)$$
$$\times\int \Big(L(\boldsymbol{z}_1,\boldsymbol{z}_2)\omega_{3,2}(\boldsymbol{z}_0,\boldsymbol{z}_1,\boldsymbol{z}_2) - \widetilde{\Gamma}_2(\boldsymbol{z}_0,\boldsymbol{z}_1)\Big)^2 p_B(\boldsymbol{z}_0,\boldsymbol{z}_1,\boldsymbol{z}_2)d\boldsymbol{z}_2$$
$$= p_B(\boldsymbol{z}_0,\boldsymbol{z}_1)Q_1^2(\boldsymbol{z}_0,\boldsymbol{z}_1)$$
$$+ p_B(\boldsymbol{z}_0,\boldsymbol{z}_1)Q_2^2(\boldsymbol{z}_0,\boldsymbol{z}_1)\times\int \Big(L(\boldsymbol{z}_1,\boldsymbol{z}_2)\omega_{3,2}(\boldsymbol{z}_0,\boldsymbol{z}_1,\boldsymbol{z}_2) - \widetilde{\Gamma}_2(\boldsymbol{z}_0,\boldsymbol{z}_1)\Big)^2$$
$$\times \tilde{f}(\overrightarrow{\boldsymbol{z}_1\boldsymbol{z}_2},\overrightarrow{\boldsymbol{z}_0\boldsymbol{z}_1},\boldsymbol{z}_1)\left|(\boldsymbol{n}(\boldsymbol{z}_1)\cdot\overrightarrow{\boldsymbol{z}_2\boldsymbol{z}_1})\right| s(\boldsymbol{z}_1,\boldsymbol{z}_2)d\boldsymbol{z}_2 d\boldsymbol{z}_2$$

and

$$\langle\langle C\rangle_B^2\rangle_F - \langle\langle C\rangle\rangle^2$$
$$= E^2(\boldsymbol{z}_0)\int \Big(L(\boldsymbol{z}_0,\boldsymbol{z}_1)w_{3,1+2}(\boldsymbol{z}_0,\boldsymbol{z}_1) - \widetilde{\Phi}\Big)^2 p_B(\boldsymbol{z}_0,\boldsymbol{z}_1)d\boldsymbol{z}_1$$
$$+\int E^2(\boldsymbol{z}_0,\boldsymbol{z}_1)w_{3,1+2}^2(\boldsymbol{z}_0,\boldsymbol{z}_1)\Big(L(\boldsymbol{z}_1,\boldsymbol{z}_2)\omega_{3,2}(\boldsymbol{z}_0,\boldsymbol{z}_1,\boldsymbol{z}_2) - \widetilde{\Gamma}_2(\boldsymbol{z}_0,\boldsymbol{z}_1)\Big)^2$$
$$\times \tilde{f}(\overrightarrow{\boldsymbol{z}_1\boldsymbol{z}_2},\overrightarrow{\boldsymbol{z}_0\boldsymbol{z}_1},\boldsymbol{z}_1)\left|(\boldsymbol{n}(\boldsymbol{z}_1)\cdot\overrightarrow{\boldsymbol{z}_2\boldsymbol{z}_1})\right| s(\boldsymbol{z}_1,\boldsymbol{z}_2)p_B(\boldsymbol{z}_1,\boldsymbol{z}_2)d\boldsymbol{z}_1 d\boldsymbol{z}_2$$

Notice the integration $\int \cdots d\boldsymbol{z}_1$ follows the same conventions as $\int \cdots d\boldsymbol{z}_2$ described above (Fig. 1). That is, it runs over the ends of all BMCRT rays emitted from $\boldsymbol{z}_0$. If this ray does not hit any scene surface, it just means it hits a large black sphere enclosing the scene) where luminance and BDF vanish.

3.3 Full noise

Substituting the $\langle\langle C^2\rangle\rangle$, $\langle\langle C\rangle_B^2\rangle_F$ and $\langle(\langle C\rangle_F - \langle\langle C\rangle\rangle)^2\rangle_B$ into (3) we obtain the following expression for the full noise in the given pixel:

$$
\begin{aligned}
V \;\approx\;& \frac{1}{n_F}\int w_{3,0}^2(\boldsymbol{z}_0,\boldsymbol{z}_1) f^2(\overrightarrow{\boldsymbol{z}_1\boldsymbol{z}_0},\overrightarrow{\boldsymbol{y}_{-1}\boldsymbol{z}_0},\boldsymbol{z}_0) L(\boldsymbol{z}_0,\boldsymbol{z}_1)\left|(\boldsymbol{n}(\boldsymbol{z}_0)\cdot\overrightarrow{\boldsymbol{z}_1\boldsymbol{z}_0})\right| s(\boldsymbol{z}_0,\boldsymbol{z}_1) d\boldsymbol{z}_1 \\
+\;& \frac{E(\boldsymbol{z}_0)}{N_B n_F}\int w_{3,1+2}^2(\boldsymbol{z}_0,\boldsymbol{z}_1)\omega_{3,1}^2(\boldsymbol{z}_0,\boldsymbol{z}_1,\boldsymbol{z}_2) f(\overrightarrow{\boldsymbol{z}_2\boldsymbol{z}_1},\overrightarrow{\boldsymbol{z}_0\boldsymbol{z}_1},\boldsymbol{z}_1)\rho(\boldsymbol{z}_0,\boldsymbol{z}_1,\boldsymbol{z}_2) d\boldsymbol{z}_1 d\boldsymbol{z}_2 \\
+\;& \frac{1}{N_B n_F}\int w_{3,1+2}^2(\boldsymbol{z}_0,\boldsymbol{z}_1)\omega_{3,2}^2(\boldsymbol{z}_0,\boldsymbol{z}_1,\boldsymbol{z}_2) E(\boldsymbol{z}_0,\boldsymbol{z}_1)\frac{b(\boldsymbol{z}_1,\boldsymbol{z}_2)}{L(\boldsymbol{z}_1,\boldsymbol{z}_2)}\rho(\boldsymbol{z}_0,\boldsymbol{z}_1,\boldsymbol{z}_2) d\boldsymbol{z}_1 d\boldsymbol{z}_2 \\
+\;& \frac{1}{N_B}\int E^2(\boldsymbol{z}_0,\boldsymbol{z}_1) w_{3,1+2}^2(\boldsymbol{z}_0,\boldsymbol{z}_1)\left(L(\boldsymbol{z}_1,\boldsymbol{z}_2)\omega_{3,2}(\boldsymbol{z}_0,\boldsymbol{z}_1,\boldsymbol{z}_2) - \widetilde{\Gamma}_2(\boldsymbol{z}_0,\boldsymbol{z}_1)\right)^2 \\
& \qquad \times\tilde{f}(\overrightarrow{\boldsymbol{z}_2\boldsymbol{z}_1},\overrightarrow{\boldsymbol{z}_0\boldsymbol{z}_1},\boldsymbol{z}_1)\left|(\overrightarrow{\boldsymbol{z}_2\boldsymbol{z}_1}\cdot\boldsymbol{n}(\boldsymbol{z}_1))\right| s(\boldsymbol{z}_1,\boldsymbol{z}_2) p_B(\boldsymbol{z}_0,\boldsymbol{z}_1) d\boldsymbol{z}_1 d\boldsymbol{z}_2 \\
+\;& \frac{E^2(\boldsymbol{z}_0)}{N_B}\int \left(L(\boldsymbol{z}_0,\boldsymbol{z}_1) w_{3,1+2}(\boldsymbol{z}_0,\boldsymbol{z}_1) - \widetilde{\Phi}\right)^2 p_B(\boldsymbol{z}_0,\boldsymbol{z}_1) d\boldsymbol{z}_1 \\
-\;& \frac{1}{N_F}\langle\langle C\rangle\rangle^2
\end{aligned}
$$

where $n_F \equiv S N_F$ and we replaced $1 - N_F^{-1}$ with 1 because usually this number of forward rays is large. Notice the term $\frac{1}{N_F}\langle\langle C\rangle\rangle^2$ is an excess accuracy because we all the same had dropped terms which are $O(\frac{1}{N_F})$ and retained only $O(\frac{1}{SN_F})$ while deriving the $\langle\langle C^2\rangle\rangle$, $\langle\langle C\rangle_B^2\rangle_F$, see [15]. So later we drop it.

Collecting the terms containing $\omega_{3,m}^2$ and recalling that (eq. (40) of [15])

$$
p_B(\boldsymbol{z}_0,\boldsymbol{z}_1) = \frac{1}{E(\boldsymbol{z}_0)} f(\overrightarrow{\boldsymbol{z}_1\boldsymbol{z}_0},\overrightarrow{\boldsymbol{z}_0\boldsymbol{y}_{-1}},\boldsymbol{z}_0)\left|(\boldsymbol{n}(\boldsymbol{z}_0)\cdot\overrightarrow{\boldsymbol{z}_0\boldsymbol{z}_1})\right| s(\boldsymbol{z}_0,\boldsymbol{z}_1)
$$

we can write V as

$$
\begin{aligned}
V \;\approx\;& n_F^{-1}E(\boldsymbol{z}_0)\int w_{3,0}^2(\boldsymbol{z}_0,\boldsymbol{z}_1) f(\overrightarrow{\boldsymbol{z}_1\boldsymbol{z}_0},\overrightarrow{\boldsymbol{y}_{-1}\boldsymbol{z}_0},\boldsymbol{z}_0) L(\boldsymbol{z}_0,\boldsymbol{z}_1) p_B(\boldsymbol{z}_0,\boldsymbol{z}_1) d\boldsymbol{z}_1 \\
& + N_B^{-1}E^2(\boldsymbol{z}_0)\int \left(L(\boldsymbol{z}_0,\boldsymbol{z}_1) w_{3,1+2}(\boldsymbol{z}_0,\boldsymbol{z}_1) - \widetilde{\Phi}\right)^2 p_B(\boldsymbol{z}_0,\boldsymbol{z}_1) d\boldsymbol{z}_1 \\
& + N_B^{-1}\int w_{3,1+2}^2(\boldsymbol{z}_0,\boldsymbol{z}_1)\mathcal{V}(\boldsymbol{z}_0,\boldsymbol{z}_1) p_B(\boldsymbol{z}_0,\boldsymbol{z}_1) d\boldsymbol{z}_1 \qquad (16)
\end{aligned}
$$

where

$$
\begin{aligned}
\mathcal{V}(\boldsymbol{z}_0,\boldsymbol{z}_1) \;\equiv\; & E^2(\boldsymbol{z}_0,\boldsymbol{z}_1)\int\Big(L(\boldsymbol{z}_1,\boldsymbol{z}_2)\omega_{3,2}(\boldsymbol{z}_0,\boldsymbol{z}_1,\boldsymbol{z}_2)-\widetilde{\Gamma}_2(\boldsymbol{z}_0,\boldsymbol{z}_1)\Big)^2 \\
& \times\tilde{f}(\overrightarrow{\boldsymbol{z}_2\boldsymbol{z}_1},\overrightarrow{\boldsymbol{z}_0\boldsymbol{z}_1},\boldsymbol{z}_1)\left|(\overrightarrow{\boldsymbol{z}_2\boldsymbol{z}_1}\cdot\boldsymbol{n}(\boldsymbol{z}_1))\right|s(\boldsymbol{z}_1,\boldsymbol{z}_2)d\boldsymbol{z}_2 \\
+\; & n_F^{-1}E^2(\boldsymbol{z}_0,\boldsymbol{z}_1) \\
& \times\int\left\{\omega_{3,1}^2(\boldsymbol{z}_0,\boldsymbol{z}_1,\boldsymbol{z}_2)\tilde{f}(\overrightarrow{\boldsymbol{z}_2\boldsymbol{z}_1},\overrightarrow{\boldsymbol{z}_0\boldsymbol{z}_1},\boldsymbol{z}_1)+\omega_{3,2}^2(\boldsymbol{z}_0,\boldsymbol{z}_1,\boldsymbol{z}_2)\frac{b(\boldsymbol{z}_1,\boldsymbol{z}_2)}{L(\boldsymbol{z}_1,\boldsymbol{z}_2)}\right\} \\
& \times L(\boldsymbol{z}_1,\boldsymbol{z}_2)\tilde{f}(\overrightarrow{\boldsymbol{z}_2\boldsymbol{z}_1},\overrightarrow{\boldsymbol{z}_0\boldsymbol{z}_1},\boldsymbol{z}_1)(\boldsymbol{n}(\boldsymbol{z}_1)\cdot\overrightarrow{\boldsymbol{z}_2\boldsymbol{z}_1})s(\boldsymbol{z}_1,\boldsymbol{z}_2)d\boldsymbol{z}_2 \qquad (17)
\end{aligned}
$$

This function depends on $\omega_{3,2}(\boldsymbol{z}_0,\boldsymbol{z}_1,\boldsymbol{z}_2)$ and $\omega_{3,1}(\boldsymbol{z}_0,\boldsymbol{z}_1,\boldsymbol{z}_2) = 1 - \omega_{3,2}(\boldsymbol{z}_0,\boldsymbol{z}_1,\boldsymbol{z}_2)$ but does not depend on $w_{3,0}$ and $w_{3,1+2}$. On the contrary, the rest terms in V are independent of $\omega_{3,m}$.

4 Optimal weights

Like in [15], these are the weights which minimize V, i.e. its variation δV in response to the variation of weights is 0. Varying (16) provided that $w_{3,0} = 1 - w_{3,1+2}$ gives

$$
\begin{aligned}
\delta V \;\approx\; & -2n_F^{-1}E(\boldsymbol{z}_0)\int\delta w_{3,1+2}(\boldsymbol{z}_0,\boldsymbol{z}_1)w_{3,0}(\boldsymbol{z}_0,\boldsymbol{z}_1) \\
& \times f(\overrightarrow{\boldsymbol{z}_1\boldsymbol{z}_0},\overrightarrow{\boldsymbol{y}_{-1}\boldsymbol{z}_0},\boldsymbol{z}_0)L(\boldsymbol{z}_0,\boldsymbol{z}_1)p_B(\boldsymbol{z}_0,\boldsymbol{z}_1)d\boldsymbol{z}_1 \\
& +2N_B^{-1}E^2(\boldsymbol{z}_0)\int\delta w_{3,1+2}(\boldsymbol{z}_0,\boldsymbol{z}_1)\Big(L(\boldsymbol{z}_0,\boldsymbol{z}_1)w_{3,1+2}(\boldsymbol{z}_0,\boldsymbol{z}_1)-\widetilde{\Phi}\Big) \\
& \times L(\boldsymbol{z}_0,\boldsymbol{z}_1)p_B(\boldsymbol{z}_0,\boldsymbol{z}_1)d\boldsymbol{z}_1 \\
& -2N_B^{-1}E^2(\boldsymbol{z}_0)\delta\widetilde{\Phi}\int\Big(L(\boldsymbol{z}_0,\boldsymbol{z}_1)w_{3,1+2}(\boldsymbol{z}_0,\boldsymbol{z}_1)-\widetilde{\Phi}\Big)p_B(\boldsymbol{z}_0,\boldsymbol{z}_1)d\boldsymbol{z}_1 \\
& +2N_B^{-1}\int\delta w_{3,1+2}(\boldsymbol{z}_0,\boldsymbol{z}_1)w_{3,1+2}(\boldsymbol{z}_0,\boldsymbol{z}_1)\mathcal{V}(\boldsymbol{z}_0,\boldsymbol{z}_1)p_B(\boldsymbol{z}_0,\boldsymbol{z}_1)d\boldsymbol{z}_1 \\
& +N_B^{-1}\int w_{3,1+2}^2(\boldsymbol{z}_0,\boldsymbol{z}_1)\delta\mathcal{V}(\boldsymbol{z}_0,\boldsymbol{z}_1)p_B(\boldsymbol{z}_0,\boldsymbol{z}_1)d\boldsymbol{z}_1
\end{aligned}
$$

where $\delta\widetilde{\Phi}$ and $\delta\mathcal{V}$ are the variations of the corresponding integrals (15) and (17). Expanding $\delta\widetilde{\Phi}$ we after some simple algebra arrive at

$$\begin{aligned}
\delta V \;\approx\; & -2n_F^{-1}E(\boldsymbol{z}_0)\int \delta w_{3,1+2}(\boldsymbol{z}_0,\boldsymbol{z}_1)w_{3,0}(\boldsymbol{z}_0,\boldsymbol{z}_1)\\
& \quad \times f(\overrightarrow{\boldsymbol{z}_1\boldsymbol{z}_0},\overrightarrow{\boldsymbol{y}_{-1}\boldsymbol{z}_0},\boldsymbol{z}_0)L(\boldsymbol{z}_0,\boldsymbol{z}_1)p_B(\boldsymbol{z}_0,\boldsymbol{z}_1)d\boldsymbol{z}_1\\
& +2N_B^{-1}E^2(\boldsymbol{z}_0)\int \delta w_{3,1+2}(\boldsymbol{z}_0,\boldsymbol{z}_1)\left(L(\boldsymbol{z}_0,\boldsymbol{z}_1)w_{3,1+2}(\boldsymbol{z}_0,\boldsymbol{z}_1)-\widetilde{\Phi}\right)\\
& \quad \times L(\boldsymbol{z}_0,\boldsymbol{z}_1)p_B(\boldsymbol{z}_0,\boldsymbol{z}_1)d\boldsymbol{z}_1\\
& +2N_B^{-1}\int \delta w_{3,1+2}(\boldsymbol{z}_0,\boldsymbol{z}_1)w_{3,1+2}(\boldsymbol{z}_0,\boldsymbol{z}_1)\mathcal{V}(\boldsymbol{z}_0,\boldsymbol{z}_1)p_B(\boldsymbol{z}_0,\boldsymbol{z}_1)d\boldsymbol{z}_1\\
& +N_B^{-1}\int w_{3,1+2}^2(\boldsymbol{z}_0,\boldsymbol{z}_1)\delta\mathcal{V}(\boldsymbol{z}_0,\boldsymbol{z}_1)p_B(\boldsymbol{z}_0,\boldsymbol{z}_1)d\boldsymbol{z}_1
\end{aligned}$$

The optimal weights are those for which $\delta V = 0$ for any $\delta w_{3,1+2}$ and $\delta\omega_{3,2}$. Since $\mathcal{V}$ is independent of $w_{3,1+2}$, this happens if and only if both $\delta\mathcal{V} = 0$ and

$$\begin{aligned}
0 = & - n_F^{-1}w_{3,0}(\boldsymbol{z}_0,\boldsymbol{z}_1)f(\overrightarrow{\boldsymbol{z}_1\boldsymbol{z}_0},\overrightarrow{\boldsymbol{y}_{-1}\boldsymbol{z}_0},\boldsymbol{z}_0)L(\boldsymbol{z}_0,\boldsymbol{z}_1)\\
& + N_B^{-1}E(\boldsymbol{z}_0)\left(L(\boldsymbol{z}_0,\boldsymbol{z}_1)w_{3,1+2}(\boldsymbol{z}_0,\boldsymbol{z}_1)-\widetilde{\Phi}\right)L(\boldsymbol{z}_0,\boldsymbol{z}_1)\\
& + N_B^{-1}w_{3,1+2}(\boldsymbol{z}_0,\boldsymbol{z}_1)\frac{\mathcal{V}(\boldsymbol{z}_0,\boldsymbol{z}_1)}{E(\boldsymbol{z}_0)}
\end{aligned}$$

Recalling that $w_{3,0} = 1 - w_{3,1+2}$ it gives

$$\begin{aligned}
w_{3,1+2}(\boldsymbol{z}_0,\boldsymbol{z}_1) &= \frac{N_B f(\overrightarrow{\boldsymbol{z}_1\boldsymbol{z}_0},\overrightarrow{\boldsymbol{y}_{-1}\boldsymbol{z}_0},\boldsymbol{z}_0)+n_F E(\boldsymbol{z}_0)\widetilde{\Phi}}{N_B f(\overrightarrow{\boldsymbol{z}_1\boldsymbol{z}_0},\overrightarrow{\boldsymbol{y}_{-1}\boldsymbol{z}_0},\boldsymbol{z}_0)+E(\boldsymbol{z}_0)n_F L(\boldsymbol{z}_0,\boldsymbol{z}_1)+\frac{n_F\mathcal{V}(\boldsymbol{z}_0,\boldsymbol{z}_1)}{E(\boldsymbol{z}_0)L(\boldsymbol{z}_0,\boldsymbol{z}_1)}}\\
&= \frac{N_B \tilde{f}(\overrightarrow{\boldsymbol{z}_1\boldsymbol{z}_0},\overrightarrow{\boldsymbol{y}_{-1}\boldsymbol{z}_0},\boldsymbol{z}_0)+n_F\widetilde{\Phi}}{N_B \tilde{f}(\overrightarrow{\boldsymbol{z}_1\boldsymbol{z}_0},\overrightarrow{\boldsymbol{y}_{-1}\boldsymbol{z}_0},\boldsymbol{z}_0)+n_F L(\boldsymbol{z}_0,\boldsymbol{z}_1)+\frac{n_F\mathcal{V}(\boldsymbol{z}_0,\boldsymbol{z}_1)}{E^2(\boldsymbol{z}_0)L(\boldsymbol{z}_0,\boldsymbol{z}_1)}} \qquad (18)
\end{aligned}$$

Then, obviously

$$\begin{aligned}
\delta\mathcal{V}(\boldsymbol{z}_0,\boldsymbol{z}_1) \;=\; & 2n_F^{-1}E^2(\boldsymbol{z}_0,\boldsymbol{z}_1)\int \delta\omega_{3,2}(\boldsymbol{z}_0,\boldsymbol{z}_1,\boldsymbol{z}_2)\\
& \times\left(\alpha(\boldsymbol{z}_0,\boldsymbol{z}_1,\boldsymbol{z}_2)\omega_{3,2}(\boldsymbol{z}_0,\boldsymbol{z}_1,\boldsymbol{z}_2)-\tilde{f}(\overrightarrow{\boldsymbol{z}_2\boldsymbol{z}_1},\overrightarrow{\boldsymbol{z}_0\boldsymbol{z}_1},\boldsymbol{z}_1)-n_F\widetilde{\Gamma}_2(\boldsymbol{z}_0,\boldsymbol{z}_1)\right)\\
& \times L(\boldsymbol{z}_1,\boldsymbol{z}_2)\tilde{f}(\overrightarrow{\boldsymbol{z}_2\boldsymbol{z}_1},\overrightarrow{\boldsymbol{z}_0\boldsymbol{z}_1},\boldsymbol{z}_1)(\boldsymbol{n}(\boldsymbol{z}_1)\cdot\overrightarrow{\boldsymbol{z}_2\boldsymbol{z}_1})s(\boldsymbol{z}_1,\boldsymbol{z}_2)d\boldsymbol{z}_2
\end{aligned}$$

where

$$\alpha(\boldsymbol{z}_0,\boldsymbol{z}_1,\boldsymbol{z}_2) \;\equiv\; \tilde{f}(\overrightarrow{\boldsymbol{z}_2\boldsymbol{z}_1},\overrightarrow{\boldsymbol{z}_0\boldsymbol{z}_1},\boldsymbol{z}_1)+\frac{b(\boldsymbol{z}_1,\boldsymbol{z}_2)}{L(\boldsymbol{z}_1,\boldsymbol{z}_2)}+n_F L(\boldsymbol{z}_1,\boldsymbol{z}_2)$$

and $b(\boldsymbol{z}_1, \boldsymbol{z}_2)$ was defined in (10).

As said above, for the optimal weights $\delta\mathcal{V} = 0$ for any $\delta\omega_{3,2}$ which happens if and only if

$$0 = \alpha(\boldsymbol{z}_0, \boldsymbol{z}_1, \boldsymbol{z}_2)\omega_{3,2}(\boldsymbol{z}_0, \boldsymbol{z}_1, \boldsymbol{z}_2) - \tilde{f}(\overrightarrow{\boldsymbol{z}_2\boldsymbol{z}_1}, \overrightarrow{\boldsymbol{z}_0\boldsymbol{z}_1}, \boldsymbol{z}_1) - n_F\widetilde{\Gamma}_2(\boldsymbol{z}_0, \boldsymbol{z}_1)$$

which implies

$$\omega_{3,2}(\boldsymbol{z}_0, \boldsymbol{z}_1, \boldsymbol{z}_2) = \frac{\tilde{f}(\overrightarrow{\boldsymbol{z}_2\boldsymbol{z}_1}, \overrightarrow{\boldsymbol{z}_0\boldsymbol{z}_1}, \boldsymbol{z}_1) + n_F\widetilde{\Gamma}_2(\boldsymbol{z}_0, \boldsymbol{z}_1)}{\tilde{f}(\overrightarrow{\boldsymbol{z}_2\boldsymbol{z}_1}, \overrightarrow{\boldsymbol{z}_0\boldsymbol{z}_1}, \boldsymbol{z}_1) + \frac{b(\boldsymbol{z}_1,\boldsymbol{z}_2)}{L(\boldsymbol{z}_1,\boldsymbol{z}_2)} + n_F L(\boldsymbol{z}_1, \boldsymbol{z}_2)} \tag{19}$$

The equations (18) and (19) are of the same functional form, being integral equations of the 2nd kind, since $\widetilde{\Gamma}_2$ is the integral of $\omega_{3,2}$ and $\widetilde{\Phi}$ is the integral of $w_{3,1+2}$. We solve them exactly like in [15]. Namely, substituting $w_{3,1+2}$ from (18) into the definition of $\widetilde{\Phi}$ (15) one has

$$\begin{aligned}
\widetilde{\Phi} &= \int \frac{N_B\tilde{f}(\overrightarrow{\boldsymbol{z}_1\boldsymbol{z}_0}, \overrightarrow{\boldsymbol{y}_{-1}\boldsymbol{z}_0}, \boldsymbol{z}_0) + n_F\widetilde{\Phi}}{N_B\tilde{f}(\overrightarrow{\boldsymbol{z}_1\boldsymbol{z}_0}, \overrightarrow{\boldsymbol{y}_{-1}\boldsymbol{z}_0}, \boldsymbol{z}_0) + n_F L(\boldsymbol{z}_0, \boldsymbol{z}_1) + \frac{n_F\mathcal{V}(\boldsymbol{z}_0,\boldsymbol{z}_1)}{E^2(\boldsymbol{z}_0)L(\boldsymbol{z}_0,\boldsymbol{z}_1)}} \\
&\qquad \times L(\boldsymbol{z}_0, \boldsymbol{z}_1)\tilde{f}(\overrightarrow{\boldsymbol{z}_0\boldsymbol{z}_1}, \overrightarrow{\boldsymbol{y}_{-1}\boldsymbol{z}_0}, \boldsymbol{z}_0)\left|(\boldsymbol{n}(\boldsymbol{z}_0)\cdot\overrightarrow{\boldsymbol{z}_1\boldsymbol{z}_0})\right| s(\boldsymbol{z}_0, \boldsymbol{z}_1)d\boldsymbol{z}_1 \\
&= \int \frac{N_B\tilde{f}(\overrightarrow{\boldsymbol{z}_1\boldsymbol{z}_0}, \overrightarrow{\boldsymbol{y}_{-1}\boldsymbol{z}_0}, \boldsymbol{z}_0)}{N_B\tilde{f}(\overrightarrow{\boldsymbol{z}_1\boldsymbol{z}_0}, \overrightarrow{\boldsymbol{y}_{-1}\boldsymbol{z}_0}, \boldsymbol{z}_0) + n_F L(\boldsymbol{z}_0, \boldsymbol{z}_1) + \frac{n_F\mathcal{V}(\boldsymbol{z}_0,\boldsymbol{z}_1)}{E^2(\boldsymbol{z}_0)L(\boldsymbol{z}_0,\boldsymbol{z}_1)}} \\
&\qquad \times L(\boldsymbol{z}_0, \boldsymbol{z}_1)\tilde{f}(\overrightarrow{\boldsymbol{z}_0\boldsymbol{z}_1}, \overrightarrow{\boldsymbol{y}_{-1}\boldsymbol{z}_0}, \boldsymbol{z}_0)\left|(\boldsymbol{n}(\boldsymbol{z}_0)\cdot\overrightarrow{\boldsymbol{z}_1\boldsymbol{z}_0})\right| s(\boldsymbol{z}_0, \boldsymbol{z}_1)d\boldsymbol{z}_1 \\
&\quad +\widetilde{\Phi}\int \frac{n_F L(\boldsymbol{z}_0, \boldsymbol{z}_1)}{N_B\tilde{f}(\overrightarrow{\boldsymbol{z}_1\boldsymbol{z}_0}, \overrightarrow{\boldsymbol{y}_{-1}\boldsymbol{z}_0}, \boldsymbol{z}_0) + n_F L(\boldsymbol{z}_0, \boldsymbol{z}_1) + \frac{n_F\mathcal{V}(\boldsymbol{z}_0,\boldsymbol{z}_1)}{E^2(\boldsymbol{z}_0)L(\boldsymbol{z}_0,\boldsymbol{z}_1)}} \\
&\qquad \times \tilde{f}(\overrightarrow{\boldsymbol{z}_0\boldsymbol{z}_1}, \overrightarrow{\boldsymbol{y}_{-1}\boldsymbol{z}_0}, \boldsymbol{z}_0)\left|(\boldsymbol{n}(\boldsymbol{z}_0)\cdot\overrightarrow{\boldsymbol{z}_1\boldsymbol{z}_0})\right| s(\boldsymbol{z}_0, \boldsymbol{z}_1)d\boldsymbol{z}_1
\end{aligned}$$

which yields

$$\widetilde{\Phi} = \frac{\int \frac{N_B\tilde{f}^2(\overrightarrow{\boldsymbol{z}_1\boldsymbol{z}_0}, \overrightarrow{\boldsymbol{y}_{-1}\boldsymbol{z}_0}, \boldsymbol{z}_0)L(\boldsymbol{z}_0,\boldsymbol{z}_1)}{N_B\tilde{f}(\overrightarrow{\boldsymbol{z}_1\boldsymbol{z}_0}, \overrightarrow{\boldsymbol{y}_{-1}\boldsymbol{z}_0}, \boldsymbol{z}_0)+n_F L(\boldsymbol{z}_0,\boldsymbol{z}_1)+\frac{n_F\mathcal{V}(\boldsymbol{z}_0,\boldsymbol{z}_1)}{E^2(\boldsymbol{z}_0)L(\boldsymbol{z}_0,\boldsymbol{z}_1)}}\left|(\boldsymbol{n}(\boldsymbol{z}_0)\cdot\overrightarrow{\boldsymbol{z}_1\boldsymbol{z}_0})\right| s(\boldsymbol{z}_0, \boldsymbol{z}_1)d\boldsymbol{z}_1}{1 - \int \frac{\tilde{f}(\overrightarrow{\boldsymbol{z}_0\boldsymbol{z}_1}, \overrightarrow{\boldsymbol{y}_{-1}\boldsymbol{z}_0}, \boldsymbol{z}_0)n_F L(\boldsymbol{z}_0,\boldsymbol{z}_1)}{N_B\tilde{f}(\overrightarrow{\boldsymbol{z}_1\boldsymbol{z}_0}, \overrightarrow{\boldsymbol{y}_{-1}\boldsymbol{z}_0}, \boldsymbol{z}_0)+n_F L(\boldsymbol{z}_0,\boldsymbol{z}_1)+\frac{n_F\mathcal{V}(\boldsymbol{z}_0,\boldsymbol{z}_1)}{E^2(\boldsymbol{z}_0)L(\boldsymbol{z}_0,\boldsymbol{z}_1)}}\left|(\boldsymbol{n}(\boldsymbol{z}_0)\cdot\overrightarrow{\boldsymbol{z}_1\boldsymbol{z}_0})\right| s(\boldsymbol{z}_0, \boldsymbol{z}_1)d\boldsymbol{z}_1} \tag{20}$$

Since $\frac{n_F L(\boldsymbol{z}_0,\boldsymbol{z}_1)}{N_B\tilde{f}(\overrightarrow{\boldsymbol{z}_1\boldsymbol{z}_0}, \overrightarrow{\boldsymbol{y}_{-1}\boldsymbol{z}_0}, \boldsymbol{z}_0)+n_F L(\boldsymbol{z}_0,\boldsymbol{z}_1)+\frac{n_F\mathcal{V}(\boldsymbol{z}_0,\boldsymbol{z}_1)}{E^2(\boldsymbol{z}_0)L(\boldsymbol{z}_0,\boldsymbol{z}_1)}} \leq 1$ then the integral in the denominator does not exceed $\int \tilde{f}(\overrightarrow{\boldsymbol{z}_2\boldsymbol{z}_1}, \overrightarrow{\boldsymbol{z}_0\boldsymbol{z}_1}, \boldsymbol{z}_1)\left|(\boldsymbol{n}(\boldsymbol{z}_1)\cdot\overrightarrow{\boldsymbol{z}_2\boldsymbol{z}_1})\right| s(\boldsymbol{z}_1, \boldsymbol{z}_2)d\boldsymbol{z}_2$ taken over all scene surfaces. In turn, this does not exceed the integral over the whole hemisphere which is 1 by definition of the normalized BDF. So the denominator is ≥ 0.

Similarly, substituting $\omega_{3,2}$ from (19) into the definition of $\widetilde{\Gamma}_2$ (14) one obtains

$$\begin{aligned}
\widetilde{\Gamma}_2 &= \int \frac{\tilde{f}(\overrightarrow{\boldsymbol{z}_2\boldsymbol{z}_1}, \overrightarrow{\boldsymbol{z}_0\boldsymbol{z}_1}, \boldsymbol{z}_1) + n_F \widetilde{\Gamma}_2(\boldsymbol{z}_1)}{\alpha(\boldsymbol{z}_0, \boldsymbol{z}_1, \boldsymbol{z}_2)} L(\boldsymbol{z}_1, \boldsymbol{z}_2) \\
&\quad \times \tilde{f}(\overrightarrow{\boldsymbol{z}_1\boldsymbol{z}_2}, \overrightarrow{\boldsymbol{z}_0\boldsymbol{z}_1}, \boldsymbol{z}_1) \left|(\boldsymbol{n}(\boldsymbol{z}_1) \cdot \overrightarrow{\boldsymbol{z}_2\boldsymbol{z}_1})\right| s(\boldsymbol{z}_1, \boldsymbol{z}_2) d\boldsymbol{z}_2 \\
&= \widetilde{\Gamma}_2(\boldsymbol{z}_1) n_F \int \frac{L(\boldsymbol{z}_1, \boldsymbol{z}_2)}{\alpha(\boldsymbol{z}_0, \boldsymbol{z}_1, \boldsymbol{z}_2)} \tilde{f}(\overrightarrow{\boldsymbol{z}_1\boldsymbol{z}_2}, \overrightarrow{\boldsymbol{z}_0\boldsymbol{z}_1}, \boldsymbol{z}_1) \left|(\boldsymbol{n}(\boldsymbol{z}_1) \cdot \overrightarrow{\boldsymbol{z}_2\boldsymbol{z}_1})\right| s(\boldsymbol{z}_1, \boldsymbol{z}_2) d\boldsymbol{z}_2 \\
&\quad + \int \frac{\tilde{f}(\overrightarrow{\boldsymbol{z}_1\boldsymbol{z}_2}, \overrightarrow{\boldsymbol{z}_0\boldsymbol{z}_1}, \boldsymbol{z}_1)}{\alpha(\boldsymbol{z}_0, \boldsymbol{z}_1, \boldsymbol{z}_2)} L(\boldsymbol{z}_1, \boldsymbol{z}_2) \tilde{f}(\overrightarrow{\boldsymbol{z}_1\boldsymbol{z}_2}, \overrightarrow{\boldsymbol{z}_0\boldsymbol{z}_1}, \boldsymbol{z}_1) \left|(\boldsymbol{n}(\boldsymbol{z}_1) \cdot \overrightarrow{\boldsymbol{z}_2\boldsymbol{z}_1})\right| s(\boldsymbol{z}_1, \boldsymbol{z}_2) d\boldsymbol{z}_2
\end{aligned}$$

which yields

$$\widetilde{\Gamma}_2 = \frac{\int \frac{\tilde{f}^2(\overrightarrow{\boldsymbol{z}_1\boldsymbol{z}_2}, \overrightarrow{\boldsymbol{z}_0\boldsymbol{z}_1}, \boldsymbol{z}_1) L(\boldsymbol{z}_1, \boldsymbol{z}_2)}{\tilde{f}(\overrightarrow{\boldsymbol{z}_2\boldsymbol{z}_1}, \overrightarrow{\boldsymbol{z}_0\boldsymbol{z}_1}, \boldsymbol{z}_1) + \frac{b(\boldsymbol{z}_1, \boldsymbol{z}_2)}{L(\boldsymbol{z}_1, \boldsymbol{z}_2)} + n_F L(\boldsymbol{z}_1, \boldsymbol{z}_2)} \left|(\boldsymbol{n}(\boldsymbol{z}_1) \cdot \overrightarrow{\boldsymbol{z}_2\boldsymbol{z}_1})\right| s(\boldsymbol{z}_1, \boldsymbol{z}_2) d\boldsymbol{z}_2}{1 - \int \frac{\tilde{f}(\overrightarrow{\boldsymbol{z}_2\boldsymbol{z}_1}, \overrightarrow{\boldsymbol{z}_0\boldsymbol{z}_1}, \boldsymbol{z}_1) n_F L(\boldsymbol{z}_1, \boldsymbol{z}_2)}{\tilde{f}(\overrightarrow{\boldsymbol{z}_2\boldsymbol{z}_1}, \overrightarrow{\boldsymbol{z}_0\boldsymbol{z}_1}, \boldsymbol{z}_1) + \frac{b(\boldsymbol{z}_1, \boldsymbol{z}_2)}{L(\boldsymbol{z}_1, \boldsymbol{z}_2)} + n_F L(\boldsymbol{z}_1, \boldsymbol{z}_2)} \left|(\boldsymbol{n}(\boldsymbol{z}_1) \cdot \overrightarrow{\boldsymbol{z}_2\boldsymbol{z}_1})\right| s(\boldsymbol{z}_1, \boldsymbol{z}_2) d\boldsymbol{z}_2} \tag{21}$$

The numerator in (21) is not negative. Indeed, since

$$\frac{n_F L(\boldsymbol{z}_1, \boldsymbol{z}_2)}{\tilde{f}(\overrightarrow{\boldsymbol{z}_2\boldsymbol{z}_1}, \overrightarrow{\boldsymbol{z}_0\boldsymbol{z}_1}, \boldsymbol{z}_1) + \frac{b(\boldsymbol{z}_1, \boldsymbol{z}_2)}{L(\boldsymbol{z}_1, \boldsymbol{z}_2)} + n_F L(\boldsymbol{z}_1, \boldsymbol{z}_2)} \leq 1,$$

the integral in the denominator does not exceed $\int \tilde{f}(\overrightarrow{\boldsymbol{z}_2\boldsymbol{z}_1}, \overrightarrow{\boldsymbol{z}_0\boldsymbol{z}_1}, \boldsymbol{z}_1) \left|(\boldsymbol{n}(\boldsymbol{z}_1) \cdot \overrightarrow{\boldsymbol{z}_2\boldsymbol{z}_1})\right| s(\boldsymbol{z}_1, \boldsymbol{z}_2) d\boldsymbol{z}_2$ taken over all scene surfaces. In turn, this does not exceed the integral over the whole hemisphere which is 1 by definition of the normalized BDF.

The formula (21) can be identically rewritten as

$$\widetilde{\Gamma}_2 = \frac{\int_{\mathfrak{S}_2} \frac{\tilde{f}^2(\overrightarrow{\boldsymbol{z}_1\boldsymbol{z}_2}, \overrightarrow{\boldsymbol{z}_0\boldsymbol{z}_1}, \boldsymbol{z}_1) L(\boldsymbol{z}_1, \boldsymbol{z}_2)}{\tilde{f}(\overrightarrow{\boldsymbol{z}_2\boldsymbol{z}_1}, \overrightarrow{\boldsymbol{z}_0\boldsymbol{z}_1}, \boldsymbol{z}_1) + c(\boldsymbol{z}_1, \boldsymbol{z}_2) + n_F L(\boldsymbol{z}_1, \boldsymbol{z}_2)} \left|(\boldsymbol{n}(\boldsymbol{z}_1) \cdot \overrightarrow{\boldsymbol{z}_2\boldsymbol{z}_1})\right| s(\boldsymbol{z}_1, \boldsymbol{z}_2) d\boldsymbol{z}_2}{1 - \chi_2 + \int_{\mathfrak{S}_2} \frac{\tilde{f}(\overrightarrow{\boldsymbol{z}_2\boldsymbol{z}_1}, \overrightarrow{\boldsymbol{z}_0\boldsymbol{z}_1}, \boldsymbol{z}_1)\left(\tilde{f}(\overrightarrow{\boldsymbol{z}_2\boldsymbol{z}_1}, \overrightarrow{\boldsymbol{z}_0\boldsymbol{z}_1}, \boldsymbol{z}_1) + c(\boldsymbol{z}_1, \boldsymbol{z}_2)\right)}{\tilde{f}(\overrightarrow{\boldsymbol{z}_2\boldsymbol{z}_1}, \overrightarrow{\boldsymbol{z}_0\boldsymbol{z}_1}, \boldsymbol{z}_1) + c(\boldsymbol{z}_1, \boldsymbol{z}_2) + n_F L(\boldsymbol{z}_1, \boldsymbol{z}_2)} \left|(\boldsymbol{n}(\boldsymbol{z}_1) \cdot \overrightarrow{\boldsymbol{z}_2\boldsymbol{z}_1})\right| s(\boldsymbol{z}_1, \boldsymbol{z}_2) d\boldsymbol{z}_2} \tag{22}$$

where

$$\chi_2 \equiv \int_{\mathfrak{S}_2} \tilde{f}(\overrightarrow{\boldsymbol{z}_2\boldsymbol{z}_1}, \overrightarrow{\boldsymbol{z}_0\boldsymbol{z}_1}, \boldsymbol{z}_1) \left|(\boldsymbol{n}(\boldsymbol{z}_1) \cdot \overrightarrow{\boldsymbol{z}_2\boldsymbol{z}_1})\right| s(\boldsymbol{z}_1, \boldsymbol{z}_2) d\boldsymbol{z}_2 \leq 1 \tag{23}$$

and $\mathfrak{S}_2$ is the set of the scene points seen from $\boldsymbol{z}_1$, i.e. points where the rays $\overrightarrow{\boldsymbol{z}_1\boldsymbol{z}_2}$ hit the scene surfaces (Fig. 1). If this area is big enough (i.e. nearly no rays $\overrightarrow{\boldsymbol{z}_1\boldsymbol{z}_2}$ miss scene surfaces) then χ_2 approaches 1. Otherwise it is < 1 and the denominator cannot vanish.

The same is the situation with the denominator in (20).

Knowing these $\widetilde{\Phi}$ and $\widetilde{\Gamma}_2$, the weights $\omega_{3,2}$ and $w_{3,1+2}$ can be calculated algebraically.

4.1 Correspondence to the case of 2 strategies

An important case is $\omega_{3,2} = 0$ when the problem reduces to the case of 2 strategies elaborated in [14]. Here the rays is merged at $\boldsymbol{z}_1$ or at $\boldsymbol{z}_0$. Equation (17) for $\omega_{3,2} = 0$ is

$$\begin{aligned} n_F\mathcal{V}(\boldsymbol{z}_0,\boldsymbol{z}_1) &= E^2(\boldsymbol{z}_0)\int L(\boldsymbol{z}_1,\boldsymbol{z}_2)f^2(\overrightarrow{\boldsymbol{z}_2\boldsymbol{z}_1},\overrightarrow{\boldsymbol{z}_0\boldsymbol{z}_1},\boldsymbol{z}_1)(\boldsymbol{n}(\boldsymbol{z}_1)\cdot\overrightarrow{\boldsymbol{z}_2\boldsymbol{z}_1})s(\boldsymbol{z}_1,\boldsymbol{z}_2)d\boldsymbol{z}_2 \\ &= E^2(\boldsymbol{z}_0)b(\boldsymbol{z}_0,\boldsymbol{z}_1). \end{aligned} \tag{24}$$

Therefore,

$$\frac{n_F\mathcal{V}(\boldsymbol{z}_0,\boldsymbol{z}_1)}{E^2(\boldsymbol{z}_0)L(\boldsymbol{z}_0,\boldsymbol{z}_1)} = \frac{b(\boldsymbol{z}_0,\boldsymbol{z}_1)}{L(\boldsymbol{z}_0,\boldsymbol{z}_1)}.$$

Substituting it into the formulae (18) and (20) we get

$$w_{3,1+2}(\boldsymbol{z}_0,\boldsymbol{z}_1) = \frac{N_B\tilde{f}(\overrightarrow{\boldsymbol{z}_1\boldsymbol{z}_0},\overrightarrow{\boldsymbol{y}_{-1}\boldsymbol{z}_0},\boldsymbol{z}_0)+n_F\widetilde{\Phi}}{N_B\tilde{f}(\overrightarrow{\boldsymbol{z}_1\boldsymbol{z}_0},\overrightarrow{\boldsymbol{y}_{-1}\boldsymbol{z}_0},\boldsymbol{z}_0)+n_FL(\boldsymbol{z}_0,\boldsymbol{z}_1)+\frac{b(\boldsymbol{z}_0,\boldsymbol{z}_1)}{L(\boldsymbol{z}_0,\boldsymbol{z}_1)}}$$

This up to notations coincides with equation (8) of [14]. The total noise is then

$$\begin{aligned} V \approx{}& \frac{E^2(\boldsymbol{z}_0)}{N_B}\int\left(L(\boldsymbol{z}_0,\boldsymbol{z}_1)w_{3,1+2}(\boldsymbol{z}_0,\boldsymbol{z}_1)-\widetilde{\Phi}\right)^2 p_B(\boldsymbol{z}_0,\boldsymbol{z}_1)d\boldsymbol{z}_1 \\ &+\frac{E^2(\boldsymbol{z}_0)}{N_Bn_F}\int\left(w_{3,0}^2(\boldsymbol{z}_0,\boldsymbol{z}_1)N_B\tilde{f}(\overrightarrow{\boldsymbol{z}_1\boldsymbol{z}_0},\overrightarrow{\boldsymbol{y}_{-1}\boldsymbol{z}_0},\boldsymbol{z}_0)+w_{3,1+2}^2(\boldsymbol{z}_0,\boldsymbol{z}_1)c(\boldsymbol{z}_0,\boldsymbol{z}_1)\right) \\ &\times L(\boldsymbol{z}_0,\boldsymbol{z}_1)p_B(\boldsymbol{z}_0,\boldsymbol{z}_1)d\boldsymbol{z}_1 \end{aligned} \tag{25}$$

5 The role of the integral terms

Calculation of the optimal weights requires several integral terms. First, we must know luminance $L(\boldsymbol{z}_1, \boldsymbol{z}_2)$ and $L(\boldsymbol{z}_0, \boldsymbol{z}_1)$. There are many methods to calculate it with MCRT. It is difficult to store directional luminance which we need. But it is easy to calculate its estimate, i.e. total luminance averaged over all directions from the given point $\boldsymbol{z}_2$, using the i-maps technique [16].

Then, we need $b(\boldsymbol{z}_1, \boldsymbol{z}_2)$ and $b(\boldsymbol{z}_0, \boldsymbol{z}_1)$. These integrals are much similar to the luminance. In the worst case the ratio $c(\boldsymbol{z}_1, \boldsymbol{z}_2) \equiv \frac{b(\boldsymbol{z}_1,\boldsymbol{z}_2)}{L(\boldsymbol{z}_1,\boldsymbol{z}_2)}$ can be estimated a priori. Indeed, for a Phong or Gaussian BDF and a Gaussian illumination this ratio is between $\max_{\boldsymbol{v}} f(\boldsymbol{v}, \overrightarrow{\boldsymbol{z}_1\boldsymbol{z}_2}, \boldsymbol{z}_2)$ and $2\max_{\boldsymbol{v}} f(\boldsymbol{v}, \overrightarrow{\boldsymbol{z}_1\boldsymbol{z}_2}, \boldsymbol{z}_2)$. For Lambert BDF it is equal to the maximum of BDF for any distribution of illumination. So a rough yet simple estimate is

$$c(\boldsymbol{z}_1, \boldsymbol{z}_2) \equiv \frac{b(\boldsymbol{z}_1, \boldsymbol{z}_2)}{L(\boldsymbol{z}_1, \boldsymbol{z}_2)} = 1.5 \max_{\boldsymbol{v}} f(\boldsymbol{v}, \overrightarrow{\boldsymbol{z}_1\boldsymbol{z}_2}, \boldsymbol{z}_2)$$

This gives us an about 50% accuracy which frequently is sufficient. The situation with $b(\boldsymbol{z}_0, \boldsymbol{z}_1)$ is similar. These L and b are used to calculate more difficult integrals: $\widetilde{\Gamma}_2$, $\mathcal{V}$ and $\widetilde{\Phi}$.

For some particular scenes the integral terms can be estimated analytically. Below we present several simple examples and demonstrate that the integral terms have a principal effect on weights. So one cannot drop this terms or substitute them with some simple heuristic.

5.1 The terms $\widetilde{\Gamma}_2$ and $\widetilde{\Phi}$

While they are given by mathematically trivial expressions their calculation may require very high precision for the integral in the denominator of (21), (20) because these denominators may vanish and then changing it a bit from 0.99 to 0.999 increases the resulting $\widetilde{\Gamma}_2$ or $\widetilde{\Phi}$ tenfold. This can happen when, first, almost all rays $\boldsymbol{z}_1 \to \boldsymbol{z}_2$ for non zero $\tilde{f}(\overrightarrow{\boldsymbol{z}_2\boldsymbol{z}_1}, \overrightarrow{\boldsymbol{z}_0\boldsymbol{z}_1}, \boldsymbol{z}_1)$ hit a scene surface, that is, $\chi_2 = 1$ (it is seen from (22)). Second, $n_F L(\boldsymbol{z}_1, \boldsymbol{z}_2)$ in the hit point must be large as compared with the BDF plus c. The first condition usually requires that BDF is rather sharp. The second condition, on the contrary, usually requires the opposite because for a sharp BDF $\tilde{f}(\overrightarrow{\boldsymbol{z}_2\boldsymbol{z}_1}, \overrightarrow{\boldsymbol{z}_0\boldsymbol{z}_1}, \boldsymbol{z}_1)$ is large, while the first condition requires it be small compared to $n_F L(\boldsymbol{z}_1, \boldsymbol{z}_2)$.

Since the light source flux in our derivation is 1, luminance cannot be very large over a finite area, thus the condition $n_F L(\boldsymbol{z}_1, \boldsymbol{z}_2) \gg 1$ requires that $n_F = SN_F$ is to be very large. Usually this means a very large number of light rays N_F but for so many rays BDPM works good even without MIS at all. The situation with the denominator in (20) is much similar. Again it may vanish only when almost all rays $\boldsymbol{z}_0 \to \boldsymbol{z}_1$ for non zero $\tilde{f}(\overrightarrow{\boldsymbol{z}_1\boldsymbol{z}_0}, \overrightarrow{\boldsymbol{y}_{-1}\boldsymbol{z}_0}, \boldsymbol{z}_0)$ hit a scene surface. Besides that the $n_F L(\boldsymbol{z}_0, \boldsymbol{z}_1)$ in the hit point must be large as compared with the N_B times BDF plus $\frac{n_F \mathcal{V}(\boldsymbol{z}_0,\boldsymbol{z}_1)}{E^2(\boldsymbol{z}_0)L(\boldsymbol{z}_0,\boldsymbol{z}_1)}$). Since the number of camera rays N_B is usually rather large, it is difficult to satisfy this condition. Therefore, it is rather rare that the denominator in (20) vanishes.

5.2 The term $\mathcal{V}$

Notice that it includes $\widetilde{\Gamma}_2$ so we must know the latter too.

The optimal weight $\omega_{3,2}$ minimizes the value of $\mathcal{V}$. Therefore, $\mathcal{V}$ for the optimal weight is less or equal its value for $\omega_{3,2} = 0$ given by (24):

$$0 \leq \mathcal{V}(\boldsymbol{z}_0, \boldsymbol{z}_1) \leq n_F^{-1} E^2(\boldsymbol{z}_0) b(\boldsymbol{z}_0, \boldsymbol{z}_1) \tag{26}$$

For $\omega_{3,2} = 1$ one cannot derive a general estimate. However, in case when the luminance $L(\boldsymbol{z}_1, \boldsymbol{z}_2)$ and $b(\boldsymbol{z}_1, \boldsymbol{z}_2)$ change slowly as compared to BDF and when whole BDF lobe is projected onto a surface $\boldsymbol{z}_2$, we can do the following. First, substitude $\omega_{3,2} = 1$ into (17). Then we must substitute $\widetilde{\Gamma}_2(\boldsymbol{z}_0, \boldsymbol{z}_1)$ calculated for this weight $\omega_{3,2} = 1$ i.e.

$$\widetilde{\Gamma}_2(\boldsymbol{z}_0, \boldsymbol{z}_1) = \int L(\boldsymbol{z}_1, \boldsymbol{z}_2) \tilde{f}(\overrightarrow{\boldsymbol{z}_2\boldsymbol{z}_1}, \overrightarrow{\boldsymbol{z}_0\boldsymbol{z}_1}, \boldsymbol{z}_1) \left| (\overrightarrow{\boldsymbol{z}_2\boldsymbol{z}_1} \cdot \boldsymbol{n}(\boldsymbol{z}_1)) \right| s(\boldsymbol{z}_1, \boldsymbol{z}_2) d\boldsymbol{z}_2$$

If $L(\boldsymbol{z}_1, \boldsymbol{z}_2)$ changes slowly as compared to BDF and whole BDF lobe projects onto a scene surface, we can approximately take its value for the centre of the BDF lobe $\boldsymbol{z}_2^*$: $\widetilde{\Gamma}_2(\boldsymbol{z}_0, \boldsymbol{z}_1) = L(\boldsymbol{z}_1, \boldsymbol{z}_2^*)$. We get $L(\boldsymbol{z}_1, \boldsymbol{z}_2) - \widetilde{\Gamma}_2(\boldsymbol{z}_0, \boldsymbol{z}_1) \approx 0$, so

$$\begin{aligned}
\mathcal{V}(\boldsymbol{z}_0, \boldsymbol{z}_1) &\approx n_F^{-1} E^2(\boldsymbol{z}_0, \boldsymbol{z}_1) b(\boldsymbol{z}_1, \boldsymbol{z}_2^*) \int \tilde{f}(\overrightarrow{\boldsymbol{z}_2\boldsymbol{z}_1}, \overrightarrow{\boldsymbol{z}_0\boldsymbol{z}_1}, \boldsymbol{z}_1)(\boldsymbol{n}(\boldsymbol{z}_1) \cdot \overrightarrow{\boldsymbol{z}_2\boldsymbol{z}_1}) s(\boldsymbol{z}_1, \boldsymbol{z}_2) d\boldsymbol{z}_2 \\
&= n_F^{-1} E^2(\boldsymbol{z}_0, \boldsymbol{z}_1) b(\boldsymbol{z}_1, \boldsymbol{z}_2^*)
\end{aligned}$$

and finally

$$\mathcal{V}(\boldsymbol{z}_0, \boldsymbol{z}_1) \approx n_F^{-1} E^2(\boldsymbol{z}_0, \boldsymbol{z}_1) b(\boldsymbol{z}_1, \boldsymbol{z}_2^*) \tag{27}$$

It is of the same form as (24), only b is taken for another surface. On the other hand, in both cases this is b for the surface where the rays are joined. If the optimal weight is not $\omega_{3,2} = 1$ the above is the upper bound.

Finally, the expression for $\mathcal{V}$ is sophisticated and we do not know a helpful estimate. We know the range (26). Its upper bound is exact for $\omega_{3,2} = 0$, but as a result it has little sense if the optimal weight is $\omega_{3,2} \approx 1$. For the latter case and in case the luminance and b change slower than BDF, we have (27). However these conditions are very difficult to check for an arbitrary scene. And if they are not satisfied than (27) is not an approximation for $\omega_{3,2} = 1$. As a result, even the upper bound in case the optimal weight is different. In other words,

$$n_F^{-1} E^2(\boldsymbol{z}_0, \boldsymbol{z}_1) \min\left(b(\boldsymbol{z}_1, \boldsymbol{z}_2^*), b(\boldsymbol{z}_0, \boldsymbol{z}_1)\right)$$

gives us just a scale.

5.3 Example 1: dark and well lit BDF at $\boldsymbol{z}_2$

This situation relates to small $c(\boldsymbol{z}_1, \boldsymbol{z}_2)$ which is about the maximum of BDF according to estimation above. This is why we say "BDF is dark". Then $n_F L(\boldsymbol{z}_1, \boldsymbol{z}_2)$ is usually also small but much greater than $c(\boldsymbol{z}_1, \boldsymbol{z}_2)$.

Intuitively, in this case it is better to use the 3rd strategy than the 2nd, i.e. to collect FMCRT rays at $\boldsymbol{z}_2$ than at $\boldsymbol{z}_1$. This is because the dark BDF kills many FMCRT rays, so few of them reach $\boldsymbol{z}_1$ and the noise in luminance estimation must be high.

We can use results from Appendix A.3 for piecewise constant $L(\boldsymbol{z}_1, \boldsymbol{z}_2)$ setting $L_{max} = L_{min} = L$ and $c \to 0$. Then

$$\begin{aligned}
\omega_{3,2} &\approx 1 \\
\mathcal{V}(\boldsymbol{z}_0, \boldsymbol{z}_1) &\approx E^2(\boldsymbol{z}_0, \boldsymbol{z}_1) c n_F^{-1} L \\
\frac{n_F \mathcal{V}(\boldsymbol{z}_0, \boldsymbol{z}_1)}{E^2(\boldsymbol{z}_0) L(\boldsymbol{z}_0, \boldsymbol{z}_1)} &\approx \frac{E(\boldsymbol{z}_0, \boldsymbol{z}_1)}{E(\boldsymbol{z}_0)} c
\end{aligned}$$

That is, we can just drop the term $\frac{n_F \mathcal{V}(\boldsymbol{z}_0, \boldsymbol{z}_1)}{E^2(\boldsymbol{z}_0) L(\boldsymbol{z}_0, \boldsymbol{z}_1)}$ because in our example $c \to 0$ and it is negligible. Therefore now the formula for the weight $w_{3,1+2}$ (18) becomes

$$w_{3,1+2}(\boldsymbol{z}_0, \boldsymbol{z}_1) \approx \frac{N_B \tilde{f}(\overrightarrow{\boldsymbol{z}_1 \boldsymbol{z}_0}, \overrightarrow{\boldsymbol{y_{-1}} \boldsymbol{z}_0}, \boldsymbol{z}_0) + n_F \widetilde{\Phi}}{N_B \tilde{f}(\overrightarrow{\boldsymbol{z}_1 \boldsymbol{z}_0}, \overrightarrow{\boldsymbol{y_{-1}} \boldsymbol{z}_0}, \boldsymbol{z}_0) + n_F L(\boldsymbol{z}_0, \boldsymbol{z}_1)}$$

Be this so for more general case, this would much simplify the calculations.

5.4 Example 2: $L(\boldsymbol{z}_1, \boldsymbol{z}_2)$ and $b(\boldsymbol{z}_1, \boldsymbol{z}_2)$ change slower than sharp BDF at $\boldsymbol{z}_1$

Now let $c(\boldsymbol{z}_1, \boldsymbol{z}_2)$ be not so small. Intuitively, in this case it is better to use the 3rd strategy than the 2nd because the BDF at $\boldsymbol{z}_1$ is sharp, so collecting FMCRT rays coming from different directions gives highly varying ray contributions and thus the noise will be high.

In this situation $L(\boldsymbol{z}_1, \boldsymbol{z}_2)$ and $c(\boldsymbol{z}_1, \boldsymbol{z}_2)$ are assumed to be nearly constant within the projection of the BDF lobe onto the surface 2. We can thus use the results from Appendix A.3 for piecewise constant $L(\boldsymbol{z}_1, \boldsymbol{z}_2)$ setting $L_{max} = L_{min} = L$.

BDF at $\boldsymbol{z}_1$ is assumed Gaussian of the form

$$\tilde{f}(\boldsymbol{v}, \boldsymbol{u}, \boldsymbol{z}_1) = \frac{1}{\pi w^2} e^{-\left(\frac{\vartheta}{w}\right)^2}$$

where ϑ is the angle between the outgoing ray and mirror reflection of the incident ray, w is BDF width. For a sharp BDF while not too bright illumination, when $w n_F L \to 0$, the equations for $\omega_{3,2}$, $\mathcal{V}(\boldsymbol{z}_0, \boldsymbol{z}_1)$, $\frac{n_F \mathcal{V}(\boldsymbol{z}_0, \boldsymbol{z}_1)}{E^2(\boldsymbol{z}_0) L(\boldsymbol{z}_0, \boldsymbol{z}_1)}$ are the same as for the example 1 (Section 5.3).

In the opposite case, when $w n_F L$ is large we have

$$\begin{aligned}
\omega_{3,2} &\approx \frac{1}{1 + 2\pi w^2 c} = \begin{cases} 1, & c \to 0 \\ 0, & c \to \infty \end{cases} \\
\mathcal{V}(\boldsymbol{z}_0, \boldsymbol{z}_1) &\approx E^2(\boldsymbol{z}_0, \boldsymbol{z}_1) \frac{c}{2\pi w^2 c + 1} n_F^{-1} L \\
\frac{n_F \mathcal{V}(\boldsymbol{z}_0, \boldsymbol{z}_1)}{E^2(\boldsymbol{z}_0) L(\boldsymbol{z}_0, \boldsymbol{z}_1)} &\approx \frac{E(\boldsymbol{z}_0, \boldsymbol{z}_1)}{E(\boldsymbol{z}_0)} \frac{c}{2\pi w^2 c + 1} = \frac{E(\boldsymbol{z}_0, \boldsymbol{z}_1)}{E(\boldsymbol{z}_0)} \begin{cases} c, & c \to 0 \\ \frac{1}{2\pi w^2}, & c \to \infty \end{cases}
\end{aligned}$$

Notice that for our Gaussian BDF and constant illumination throughout the projection of the BDF lobe the function $c(\boldsymbol{z}_0, \boldsymbol{z}_1)$ (defined similarly to $c(\boldsymbol{z}_1, \boldsymbol{z}_2)$ but for the arguments $\boldsymbol{z}_0, \boldsymbol{z}_1$ instead of $\boldsymbol{z}_1, \boldsymbol{z}_2$) for this (1st) surface is equal to $\frac{1}{2\pi w^2}$. This naturally agrees with (24).

Therefore, for both $\omega_{3,2} \approx 0$ and $\omega_{3,2} \approx 1$ the value of $\frac{n_F \mathcal{V}(\boldsymbol{z}_0, \boldsymbol{z}_1)}{E^2(\boldsymbol{z}_0) L(\boldsymbol{z}_0, \boldsymbol{z}_1)}$ is close to the scaled c if we take c for the proper surface:

$$\frac{n_F \mathcal{V}(\boldsymbol{z}_0, \boldsymbol{z}_1)}{E^2(\boldsymbol{z}_0) L(\boldsymbol{z}_0, \boldsymbol{z}_1)} \approx \begin{cases} \frac{E(\boldsymbol{z}_0, \boldsymbol{z}_1)}{E(\boldsymbol{z}_0)} c(\boldsymbol{z}_1, \boldsymbol{z}_2), & \omega_{3,2} \approx 1 \\ c(\boldsymbol{z}_0, \boldsymbol{z}_1), & \omega_{3,2} \approx 0 \end{cases} \tag{28}$$

The argument $\boldsymbol{z}_2$ is absent in the left hand side, but since in this example $c(\boldsymbol{z}_1, \boldsymbol{z}_2) = const$ we may use any $\boldsymbol{z}_2$. Notice the scale factor $\frac{E(\boldsymbol{z}_0, \boldsymbol{z}_1)}{E(\boldsymbol{z}_0)}$, i.e. the integral of the BDF at $\boldsymbol{z}_1$, naturally appears only when the rays merge at $\boldsymbol{z}_2$ after having been scattered at $\boldsymbol{z}_1$. If the rays are merged at $\boldsymbol{z}_1$, the BDF at $\boldsymbol{z}_1$ did not scatter them, so this scale factor is naturally absent.

Therefore now the formula for the weight $w_{3,1+2}$ (18) becomes

$$w_{3,1+2}(\boldsymbol{z}_0,\boldsymbol{z}_1) = \begin{cases} \frac{N_B \tilde{f}(\overrightarrow{\boldsymbol{z}_1\boldsymbol{z}_0},\overrightarrow{\boldsymbol{y}_{-1}\boldsymbol{z}_0},\boldsymbol{z}_0)+n_F\widetilde{\Phi}}{N_B \tilde{f}(\overrightarrow{\boldsymbol{z}_1\boldsymbol{z}_0},\overrightarrow{\boldsymbol{y}_{-1}\boldsymbol{z}_0},\boldsymbol{z}_0)+n_F L(\boldsymbol{z}_0,\boldsymbol{z}_1)+\frac{E(\boldsymbol{z}_0,\boldsymbol{z}_1)}{E(\boldsymbol{z}_0)}c(\boldsymbol{z}_1,\boldsymbol{z}_2)}, & \omega_{3,2}\approx 1 \\ \frac{N_B \tilde{f}(\overrightarrow{\boldsymbol{z}_1\boldsymbol{z}_0},\overrightarrow{\boldsymbol{y}_{-1}\boldsymbol{z}_0},\boldsymbol{z}_0)+n_F\widetilde{\Phi}}{N_B \tilde{f}(\overrightarrow{\boldsymbol{z}_1\boldsymbol{z}_0},\overrightarrow{\boldsymbol{y}_{-1}\boldsymbol{z}_0},\boldsymbol{z}_0)+n_F L(\boldsymbol{z}_0,\boldsymbol{z}_1)+c(\boldsymbol{z}_0,\boldsymbol{z}_1)}, & \omega_{3,2}\approx 0 \end{cases} \tag{29}$$

The case $\omega_{3,2} \approx 0$ naturally results in exactly the same formula as in the case of two strategies.

Again, the point $\boldsymbol{z}_2$ which we need in case $\omega_{3,2} \approx 1$ is absent in the left hand side (the separable weight can *not* depend on it). We can substitute any point in our example, when $c(\boldsymbol{z}_1,\boldsymbol{z}_2) = const$, but what to do in more general case, when $c(\boldsymbol{z}_1,\boldsymbol{z}_2)$ does vary, at least slowly? A possible solution is to take $\boldsymbol{z}_2 = \boldsymbol{z}_2^*$ where $\boldsymbol{z}_2^*$ is the hit point of the specular scattering of the ray $\overrightarrow{\boldsymbol{z}_0\boldsymbol{z}_1}$. This is natural if the BDF at $\boldsymbol{z}_1$ is very sharp (glossy), but what if it is not?

Before bothering how to resolve this problem, let us first see if this is worth effort at all, i.e. if the above decision (to use c for the surface chosen depending on $\omega_{3,2}$) works when $c(\boldsymbol{z}_1,\boldsymbol{z}_2) = const$ holds while the weight $\omega_{3,2}$ depends on $\boldsymbol{z}_2$.

In the current example $\omega_{3,2}$ depends on the relation between the brightness of illumination $L(\boldsymbol{z}_1,\boldsymbol{z}_2)$ and the BDF width w. Currently $L(\boldsymbol{z}_1,\boldsymbol{z}_2)$ was spatially uniform, so a natural decision is to make it spatially variable and see what happens. Let's make the situation maximally different from current uniform luminance, e.g. make it to change fast from large (maximum) to small (minimum) values.

5.5 Example 3: high-frequency $L(\boldsymbol{z}_1,\boldsymbol{z}_2)$ and sharp BDF at $\boldsymbol{z}_1$

Let the $\mathfrak{S}_2$ be a set of the uniformly distributed points which randomly take values of either L_{max} or L_{min}. This is a high-frequency $L(\boldsymbol{z}_1,\boldsymbol{z}_2)$. The points where $L(\boldsymbol{z}_1,\boldsymbol{z}_2) = L_{max}$ occupy the fraction a of the total area of $\mathfrak{S}_2$. In the rest of $\mathfrak{S}_2$ (whose area fraction is therefore $1-a$) the luminance $L(\boldsymbol{z}_1,\boldsymbol{z}_2) = L_{min}$. The BDF at $\boldsymbol{z}_1$ is assumed Gaussian like in Section (5.4).

One can understand the role of spatial variations of luminance as follows. The term $\mathcal{V}$ is nothing but the noise of the 2nd segment. Indeed, let us substitute in (25)

$$N_B \mapsto 1, \qquad w_{3,1+2} \mapsto \omega_{3,2},$$
$$\boldsymbol{z}_{-1} \mapsto \boldsymbol{z}_0, \quad \boldsymbol{z}_0 \mapsto \boldsymbol{z}_1, \quad \boldsymbol{z}_1 \mapsto \boldsymbol{z}_2$$

(then automatically $\widetilde{\Phi} \mapsto \widetilde{\Gamma}_2$). The term $E(\boldsymbol{z}_0) = \mu(\overrightarrow{\boldsymbol{y}_{-1}\boldsymbol{z}_0},\boldsymbol{z}_0)$ in (25) is the energy of camera ray before hitting $\boldsymbol{z}_1$, so it must turn into $\mu(\overrightarrow{\boldsymbol{z}_0\boldsymbol{z}_1},\boldsymbol{z}_1)$ which equals $\frac{E(\boldsymbol{z}_0,\boldsymbol{z}_1)}{E(\boldsymbol{z}_0)}$ by construction in (17). As a result, V turns into $\frac{\mathcal{V}(\boldsymbol{z}_0,\boldsymbol{z}_1)}{E^2(\boldsymbol{z}_0)}$ and this is equal to the full noise for the case of 2 strategies provided that the single camera ray for the pixel ($N_B = 1$) starts at $\boldsymbol{z}_0$ and hits $\boldsymbol{z}_1$, the segment $[\boldsymbol{z}_0, \boldsymbol{z}_1]$ is not random; at $\boldsymbol{z}_1$ the ray is scattered stochastically. Therefore, $\frac{\mathcal{V}(\boldsymbol{z}_0,\boldsymbol{z}_1)}{E^2(\boldsymbol{z}_0)}$ is the noise of luminance estimation taken at $\boldsymbol{z}_1$. This noise includes contribution from both BMCRT and FMCRT. Very roughly, if the BCMRT ray takes greatly varying luminance $L(\boldsymbol{z}_1, \boldsymbol{z}_2)$ from different points $\boldsymbol{z}_2$ then this is BMCRT noise. If it is large, this may make the 3rd strategy less advantageous (or not at all).

The derivation of $\mathcal{V}(\boldsymbol{z}_0, \boldsymbol{z}_1)$ is in Appendix A. For the sake of simplicity we shall consider only the case when $c(\boldsymbol{z}_1, \boldsymbol{z}_2)$ is constant and small. Below we reproduce only those final formulae that are to be used here. For a sharp BDF while not too bright illumination, when $w^2 n_F L_{max} \to 0$ and

$$\begin{aligned}
\omega_{3,2} &\approx 1 \\
\mathcal{V}(\boldsymbol{z}_0, \boldsymbol{z}_1) &\approx E^2(\boldsymbol{z}_0, \boldsymbol{z}_1) a(1-a)(L_{max} - L_{min})^2 \\
\frac{n_F \mathcal{V}(\boldsymbol{z}_0, \boldsymbol{z}_1)}{E^2(\boldsymbol{z}_0) L(\boldsymbol{z}_0, \boldsymbol{z}_1)} &\approx \frac{E(\boldsymbol{z}_0, \boldsymbol{z}_1)}{E(\boldsymbol{z}_0)} (1-a) n_F L_{max}
\end{aligned}$$

In the opposite case, when $w n_F L$ is large and

$$\begin{aligned}
\omega_{3,2} &\approx 0 \\
\mathcal{V}(\boldsymbol{z}_0, \boldsymbol{z}_1) &\approx E^2(\boldsymbol{z}_0, \boldsymbol{z}_1) a(1-a) \frac{1}{2\pi w^2} n_F^{-1} L_{max} \\
\frac{n_F \mathcal{V}(\boldsymbol{z}_0, \boldsymbol{z}_1)}{E^2(\boldsymbol{z}_0) L(\boldsymbol{z}_0, \boldsymbol{z}_1)} &\approx \frac{E(\boldsymbol{z}_0, \boldsymbol{z}_1)}{E(\boldsymbol{z}_0)} (1-a) \frac{1}{2\pi w^2}
\end{aligned}$$

What can one conclude from these formulae? For $\omega_{3,2} \approx 0$ it is still $\frac{n_F \mathcal{V}(\boldsymbol{z}_0,\boldsymbol{z}_1)}{E^2(\boldsymbol{z}_0) L(\boldsymbol{z}_0,\boldsymbol{z}_1)} \approx c(\boldsymbol{z}_0, \boldsymbol{z}_1)$. But for $\omega_{3,2} \approx 1$ the value of $\frac{n_F \mathcal{V}(\boldsymbol{z}_0,\boldsymbol{z}_1)}{E^2(\boldsymbol{z}_0) L(\boldsymbol{z}_0,\boldsymbol{z}_1)}$ is determined by the maximal surface luminance and not by the function $c(\boldsymbol{z}_1, \boldsymbol{z}_2)$.

Therefore, we cannot calculate the weight $w_{3,1+2}$ from a formula like (29) even though in our extreme case the weight $\omega_{3,2}$ is still nearly constant. The difference between the actual value of $\frac{n_F \mathcal{V}(\boldsymbol{z}_0,\boldsymbol{z}_1)}{E^2(\boldsymbol{z}_0) L(\boldsymbol{z}_0,\boldsymbol{z}_1)}$ and its estimation using c for the proper surface (28) can be very large. As a result, the "estimated" weight $w_{3,1+2}$ can strongly differ from its actual value. For example, for large L_{max} and small a it is possible that the actual $w_{3,1+2} \approx 0$ while the estimated one is $w_{3,1+2} \approx 1$. The less chances in a more general case. Therefore, it is impossible to calculate $w_{3,1+2}$ without actual calculation of $\mathcal{V}(\boldsymbol{z}_0, \boldsymbol{z}_1)$.

The situation with the integral terms $\widetilde{\Gamma}_2$ and $\widetilde{\Phi}$ is similar though there influence is weaker. There are rather many situations when we can drop them (set to 0) while calculating weights. But even when this is possible it does not help much because we need $\mathcal{V}$.

5.6 The integral terms: conclusions

The calculation of $\widetilde{\Gamma}_2$, $\widetilde{\Phi}$ and $\mathcal{V}$ is expensive while we do not have a due approximations. All what we have is:

- For $\widetilde{\Gamma}_2$ we have "range values" 0 and $\max_{\boldsymbol{z}_2 \in \mathfrak{S}_2} L(\boldsymbol{z}_1, \boldsymbol{z}_2)$. The value of $\widetilde{\Gamma}_2$ affects both the weight $\omega_{3,2}$ explicitly and the weight $w_{3,1+2}$ via $\mathcal{V}$.
- For $\widetilde{\Phi}$ we have "range values" 0 and $\max_{\boldsymbol{z}_1 \in \mathfrak{S}_1} L(\boldsymbol{z}_0, \boldsymbol{z}_1)$. The value of $\widetilde{\Phi}$ affects the weight $w_{3,1+2}$ explicitly.
- In case the calculated value of $\omega_{3,2}$ is close to 0, we have the case of 2 strategies and $\mathcal{V}$ is known, (24). Otherwise we have no helpful estimate. We can believe $n_F^{-1} E^2(\boldsymbol{z}_0, \boldsymbol{z}_1) b(\boldsymbol{z}_0, \boldsymbol{z}_1)$ gives a characteristic scale but no more.

As it has been demonstrated by examples (Sections (5.3)-(5.5)) we must use the exact calculated value for $\mathcal{V}$ (and correspondingly for $\widetilde{\Gamma}_2$) and cannot replace them with a helpful approximation.

6 Results

Here, we present and analyze the results of the computational experiments done for a benchmark scene. The scene consists of three surfaces. The light source illuminates the 3rd one, whose scattered light illuminates the 2nd surface and then the 1st surface. The 1st surface is seen by the camera (Fig. 2).

This scene is rather general as considering the first three segments of the camera paths. Its simplification is that the 3rd surface is illuminated directly while in the general situation it could be illuminated by the secondary illumination. This is however not important for us because all our formulae do not distinguish the direct and indirect components of $L(\boldsymbol{z}_1, \boldsymbol{z}_2)$.

In the experiment we consider only the ray paths for which $\boldsymbol{z}_0$ is at the surface #1, $\boldsymbol{z}_1$ is at the surface #2, and $\boldsymbol{z}_2$ is at the surface #3. This simplifies the calculations. In principle BDPM is a summation over all light paths, so we just proceed with one its component. In our scene it dominates, i.e. light paths of other types have rather small contribution.

In calculations the BDF at $\boldsymbol{z}_0$ and at $\boldsymbol{z}_2$ is Lambert model with albedo 1, and the BDF at $\boldsymbol{z}_1$ is the Phong model with glossiness parameter γ:

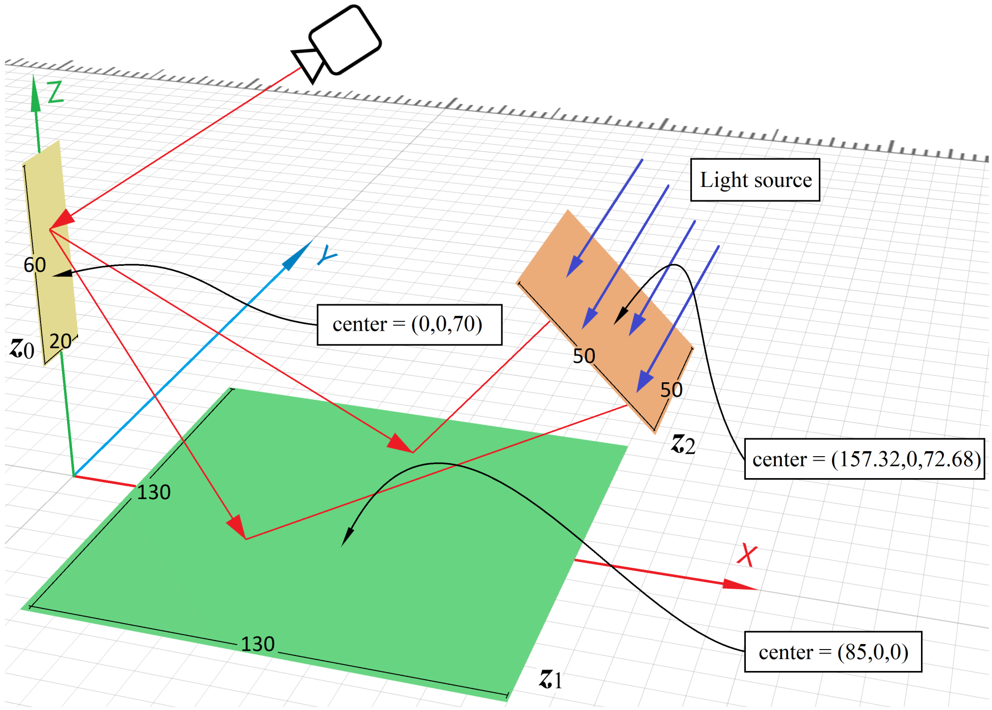

Figure 2: The benchmark scene. The yellow, green and orange rectangles are the 1st, 2nd and 3rd surfaces, respectively. The red arrows show a camera ray paths. The blue arrows show illumination from the parallel light source.

$$f(\boldsymbol{v}, \boldsymbol{u}, \boldsymbol{z}_2) \;=\; \mu_1 \frac{\gamma + 2}{2\pi} \left(\max(0, -(\boldsymbol{r} \cdot \boldsymbol{u}))\right)^{\gamma}$$

where $\boldsymbol{u}$ is the view direction, $\boldsymbol{r} = \boldsymbol{v} - 2(\boldsymbol{n} \cdot \boldsymbol{v})\boldsymbol{n}$ is the direction of the specular reflection and μ_1 is the integral reflectance $\int f_1(\boldsymbol{v}, \boldsymbol{u}) \left|(\boldsymbol{n} \cdot \boldsymbol{v})\right| d^2\boldsymbol{v}$ for normal incidence (for tangent incidence it drops), i.e. $\mu_1 = \mu(\boldsymbol{n}, \boldsymbol{z}_1)$ in terms of (8). For the Lambert BDF at $\boldsymbol{z}_2$, obviously, $c(\boldsymbol{z}_1, \boldsymbol{z}_2) = \pi^{-1}$. Light source with unit flux provides parallel illumination of the $\boldsymbol{z}_2$ domain whose size is 50×50. As a results its luminance is uniform $L(\boldsymbol{z}_1, \boldsymbol{z}_2) = \pi^{-1} \frac{1}{50 \times 50}$.

In all cases the number of camera rays per pixel is $N_B = 25$ and the radius of the integration sphere is $R = 0.3$.

The image seen by camera are very simple: it is rectangle (the 1st surface) in the perspective projection (Fig 3). Four cases are represented in Fig. 3. In first three "pure" cases (BDD=0, 1, 2) all rays are merged at the predefined vertex. The "BDD" means "backward diffuse depth", i.e. the number of diffuse scattering events the camera ray underwent before the integration sphere. BDD=0 means merging at the first surface, BDD=1 – at the second and BDD=2 – at the third ones. "MIS" is the optimal weights (mixing strategies) described in this paper. The simulation was done for a small number of BDPM iterations to make the noise well visible.

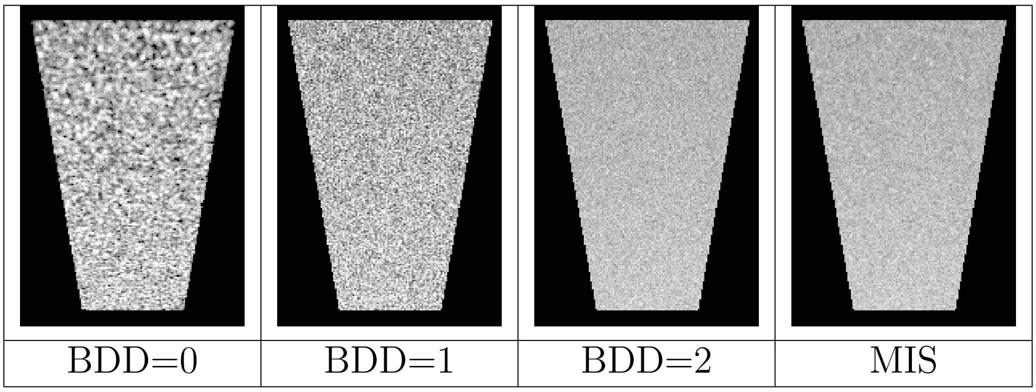

Figure 3: Camera images, i.e. the distribution of luminance $L(\boldsymbol{y_{-1}}, \boldsymbol{z}_0)$ for $\mu_1 = 0.1$, $\gamma = 900$ and $N_F = 530000$ after 100 BDPM iterations.

The tables 1 and 2 show the relative noise $\frac{\sqrt{V(\boldsymbol{y_{-1}},\boldsymbol{z}_0)}}{L(\boldsymbol{y_{-1}},\boldsymbol{z}_0)}$ in the center of the image, averaged over 10x10 pixels for different scene parameters and cases. Table 1 is for a relatively small number of light rays ($N_F = 530000$ and therefore $n_F L(\boldsymbol{z}_1, \boldsymbol{z}_2) = 19.08$) while in Table 2 this number is doubled, so we can see the role of this parameter.

Table 1: Relative noise $\frac{\sqrt{V(\boldsymbol{y_{-1}},\boldsymbol{z}_0)}}{L(\boldsymbol{y_{-1}},\boldsymbol{z}_0)}$ in the image center for $N_F = 530000$ and different parameters of the BDF at $\boldsymbol{z}_1$ after 1000 BDPM iterations.

BDF at $\boldsymbol{z}_1$		**BDD=0**	**BDD=1**	**BDD=2**	**MIS**
integral μ_1	**glossiness γ**				
0.1	**100**	459%	201%	187%	155%
	300	455%	297%	185%	166%
	600	454%	387%	186%	169%
	900	449%	464%	185%	169%
0.3	**100**	267%	204%	186%	137%
	300	266%	295%	183%	146%
	600	264%	390%	186%	149%
	900	263%	473%	190%	153%
1.0	**100**	143%	205%	182%	105%
	300	145%	294%	184%	111%
	600	143%	384%	185%	112%
	900	144%	462%	183%	112%

The results are rather intuitive. When BDF at $\boldsymbol{z}_1$ is smooth (γ is small) and the 2nd surface is bright (integral μ_1 is large), one can merge the camera and the light rays at any surface without much gain or loss in the noisiness. When this surface is dark (integral μ_1 is small), illumination of the 1st surface is low, and thus merging rays at this surface (case BDD=0) results in high relative

Table 2: Relative noise $\frac{\sqrt{V(\boldsymbol{y}_{-1},\boldsymbol{z}_0)}}{L(\boldsymbol{y}_{-1},\boldsymbol{z}_0)}$ in the image center for $N_F = 1060000$ and different parameters of the BDF at $\boldsymbol{z}_1$ after 650 BDPM iterations.

BDF at $\boldsymbol{z}_1$		**BDD=0**	**BDD=1**	**BDD=2**	**MIS**
integral μ_1	**glossiness γ**				
0.1	**100**	315%	174%	185%	140%
	300	320%	242%	187%	155%
	600	314%	298%	183%	154%
	900	316%	359%	186%	158%
0.3	**100**	185%	177%	185%	119%
	300	183%	241%	183%	124%
	600	184%	293%	184%	128%
	900	186%	346%	186%	128%
1.0	**100**	103%	176%	185%	84%
	300	102%	240%	188%	88%
	600	102%	301%	185%	89%
	900	102%	344%	185%	89%

noise. Meanwhile, since our BMCRT does not kill rays even on a dark surface, the number of camera rays reached the 2nd and the 3rd surface remains the same. Therefore the relative noise while merging rays at these surfaces does not grow.

Mixing strategies (the last column) improves the situation for any parameters. The noise for MIS is always less than for other cases.

One can see that results in Table 2 are qualitatively the same, though now the case BDD=0 is better than in Table 1 for dark BDF at $\boldsymbol{z}_1$ as well as the case BDD=1 is better for sharp BDF at $\boldsymbol{z}_1$. This is again intuitively expected because the disadvantage was due to the noise coming from gathering luminance from all FMCRT rays that hit the integration sphere. Now their number grew proportionally to the increased N_B, i.e. twice, and thus its contribution to the total noise dropped.

7 Discussion and conclusions

In [14, 15] we have developed a reduced MIS (which mixes a limited number of strategies whose weights depend on a limited number of the first ray path vertices) for BDPM. For two strategies the resulting expressions for the weights were rather compact and understandable and we developed a general numeric method applied for a simple test scene [14]. For three strategies, these expressions appear so sophisticated and complex [15] that their numeric calculation even in simple cases is expensive and cancels all the benefits of the expected better convergence of BDPM.

Meanwhile, it is not rare that the third (and even the fourth) strategy is really needed because the two ones are all bad. For example, it is the case when a few light rays achieve for the 1st camera ray vertex while at the second one BDF is sharp so its integration with a diffuse illumination is highly noisy. This case was considered even for the very simple benchmark scene (Section 6).

In current work we developed a reduced MIS which operates separable weights. Roughly speaking, it splits the problem of choosing at which of the 3 camera vertices: $\boldsymbol{z}_0$, $\boldsymbol{z}_1$ or $\boldsymbol{z}_2$, to merge the camera and light rays. First, we choose whether to merge at $\boldsymbol{z}_0$ or at any of the next vertices ($\boldsymbol{z}_1$ or $\boldsymbol{z}_2$) and determine the weights for these two choices. Then we choose whether to merge at $\boldsymbol{z}_1$ or at $\boldsymbol{z}_2$ and determine the weights for these choices. As a result, the weight of the choice "merge at $\boldsymbol{z}_2$" becomes a product of the weight of the choice "at any of the next vertices ($\boldsymbol{z}_1$ or $\boldsymbol{z}_2$)" and the weight of the choice "$\boldsymbol{z}_2$ among the latter".

Unlike the original MIS for 3 strategies [15], this approach admits a rather straightforward generalization for n strategies. Now we first choose between two strategies: whether to merge rays at $\boldsymbol{z}_0$ or at any of $\boldsymbol{z}_1$, $\boldsymbol{z}_2$, ..., $\boldsymbol{z}_{n-1}$. Then we choose whether to merge at $\boldsymbol{z}_1$ or at any of $\boldsymbol{z}_2$, $\boldsymbol{z}_3$, ..., $\boldsymbol{z}_{n-1}$, and so on. In each step we have the choice between two strategies. As a results, the final weights are the products of the "relative" weights at each step. This generalizes the formula (1):

$$\begin{aligned}
w_{n,0}(\boldsymbol{z}_0,...,\boldsymbol{z}_{n-1}) &= w_{n,0}(\boldsymbol{z}_0,\boldsymbol{z}_1)\\
w_{n,1}(\boldsymbol{z}_0,...,\boldsymbol{z}_{n-1}) &= w_{n,1+2+\cdots}(\boldsymbol{z}_0,\boldsymbol{z}_1)w_{n-1,0}(\boldsymbol{z}_0,\boldsymbol{z}_1,\boldsymbol{z}_2)\\
w_{n,2}(\boldsymbol{z}_0,...,\boldsymbol{z}_{n-1}) &= w_{n,1+2+\cdots}(\boldsymbol{z}_0,\boldsymbol{z}_1)w_{n-1,1+2+\cdots}(\boldsymbol{z}_0,\boldsymbol{z}_1,\boldsymbol{z}_2)w_{n-2,0}(\boldsymbol{z}_1,\boldsymbol{z}_2,\boldsymbol{z}_3)\\
w_{n,3}(\boldsymbol{z}_0,...,\boldsymbol{z}_{n-1}) &= w_{n,1+2+\cdots}(\boldsymbol{z}_0,\boldsymbol{z}_1)w_{n-1,1+2+\cdots}(\boldsymbol{z}_0,\boldsymbol{z}_1,\boldsymbol{z}_2)\\
&\quad\times w_{n-2,1+2+\cdots}(\boldsymbol{z}_1,\boldsymbol{z}_2,\boldsymbol{z}_3)w_{n-3,0}(\boldsymbol{z}_2,\boldsymbol{z}_3,\boldsymbol{z}_4)\\
&\cdots\cdots\cdots\\
w_{n,n-1}(\boldsymbol{z}_0,...,\boldsymbol{z}_{n-1}) &= w_{n,1+2+\cdots}(\boldsymbol{z}_0,\boldsymbol{z}_1)w_{n-1,1+2+\cdots}(\boldsymbol{z}_0,\boldsymbol{z}_1,\boldsymbol{z}_2)\\
&\quad\times w_{n-2,1+2+\cdots}(\boldsymbol{z}_1,\boldsymbol{z}_2,\boldsymbol{z}_3)w_{1,0}(\boldsymbol{z}_{n-1},\boldsymbol{z}_n,\boldsymbol{z}_{n+1})
\end{aligned} \tag{30}$$

with $w_{1,0}(\boldsymbol{z}_{n-1},\boldsymbol{z}_n,\boldsymbol{z}_{n+1})\equiv 1$ and normalization

$$w_{n-m,0}(\boldsymbol{z}_{m-1},\boldsymbol{z}_m,\boldsymbol{z}_{m+1})+w_{n-m,1+2+\cdots}(\boldsymbol{z}_{m-1},\boldsymbol{z}_m,\boldsymbol{z}_{m+1})=1,\quad m=0,1,...,n-2$$

All weights but $w_{n,0}$ and $w_{n,1+2+\cdots}$ have three arguments. As to $w_{n,0}$ and $w_{n,1+2+\cdots}$, they may be considered as also having have two ones because they implicitly depend on the camera origin $\boldsymbol{y}_{-1}$ which is the same for all rays and all pixels and thus is usually dropped from arguments. The weights in the right hand side of (30) are given by expressions similar to (18), (19).

In current work $n = 3$ and for historic reasons we used notation $\omega_{2,*}$ instead of $w_{n-1,*}$:

$$\begin{aligned} w_{n,0}(\boldsymbol{z}_0, \boldsymbol{z}_1) &= w_{3,0}(\boldsymbol{z}_0, \boldsymbol{z}_1) \\ w_{n,1+2+\cdots}(\boldsymbol{z}_0, \boldsymbol{z}_1) &= w_{3,1+2}(\boldsymbol{z}_0, \boldsymbol{z}_1) \\ w_{n-1,0}(\boldsymbol{z}_0, \boldsymbol{z}_1, \boldsymbol{z}_2) &= \omega_{2,0}(\boldsymbol{z}_0, \boldsymbol{z}_1, \boldsymbol{z}_2) \\ w_{n-1,1+2+\cdots}(\boldsymbol{z}_0, \boldsymbol{z}_1, \boldsymbol{z}_2) &= \omega_{2,1}(\boldsymbol{z}_0, \boldsymbol{z}_1, \boldsymbol{z}_2) \end{aligned}$$

We had considered the case $n = 3$ in detail and produced the final expressions for the separable weights. They are much simpler than in [15] and allowed for a numerical calculations at least for a simplified example of the benchmark scene in Section 6

But even in this case the calculations are difficult. There are integral terms, the worst of them being $\mathcal{V}$ whose integrands include the other integral terms. As to the $\widetilde{\Gamma}_2$ and $\widetilde{\Phi}$, they are the fractions whose denominators are of the form 1 minus integral, which integral can be arbitrary close to 1. Therefore we must always calculate it with high accuracy.

In Section 5 we demonstrated that these integral terms are really important. So one cannot just drop them or use some heuristic substitutes for them, but must calculate them with due accuracy. For the simple benchmark scene from section 6 it was possible because for this scene they can be calculated by deterministic quadratures. But for a general scene we can only use Monte Carlo integration, whose accuracy is also limited by the Monte Carlo noise. This is a sort of a vicious circle: for the weights to decrease noise, we must calculate those weights with MC until their noise is nearly gone.

Therefore the final conclusion is that MIS for n strategies with $n > 2$ is computationally inefficient. The efforts required to calculate the weights nullify the benefits of using them.

A Appendix. The case of high-frequency $L(\boldsymbol{z}_1, \boldsymbol{z}_2)$

Definition of a high-frequency $L(\boldsymbol{z}_1, \boldsymbol{z}_2)$ is in Section 5.5. The points where $L(\boldsymbol{z}_1, \boldsymbol{z}_2) = L_{max}$ occupy the fraction a of the total area of $\mathfrak{S}_2$. In the rest of $\mathfrak{S}_2$ (whose area fraction is $1 - a$) the luminance $L(\boldsymbol{z}_1, \boldsymbol{z}_2) = L_{min}$. We also assume $c(\boldsymbol{z}_1, \boldsymbol{z}_2)$ is constant.

Let us calculate weights and noise for this case.

To this end, we need to calculate integrals which include this high-frequency $L(\boldsymbol{z}_1, \boldsymbol{z}_2)$. First, it is obvious that

$$\int_{\mathfrak{S}_2} g(\boldsymbol{z}_2, L(\boldsymbol{z}_1, \boldsymbol{z}_2))d\boldsymbol{z}_2 \approx a\int_{\mathfrak{S}_2} g(\boldsymbol{z}_2, L_{max})d\boldsymbol{z}_2 + (1-a)\int_{\mathfrak{S}_2} g(\boldsymbol{z}_2, L_{min})d\boldsymbol{z}_2 \quad (31)$$

Then, apply this approximation to the integrals of the sort

$$\int g(\boldsymbol{z}_2, L(\boldsymbol{z}_1, \boldsymbol{z}_2))\tilde{f}(\overrightarrow{\boldsymbol{z}_2\boldsymbol{z}_1}, \overrightarrow{\boldsymbol{z}_0\boldsymbol{z}_1}, \boldsymbol{z}_1)\left|(\boldsymbol{n}(\boldsymbol{z}_1)\cdot\overrightarrow{\boldsymbol{z}_2\boldsymbol{z}_1})\right| s(\boldsymbol{z}_1, \boldsymbol{z}_2)d\boldsymbol{z}_2$$

whose integrand vanishes in the periphery of the BDF lobe. Assuming BDF is sharp or $\mathfrak{S}_2$ is large, all the projection of the BDF lobe onto the surface fits within $\mathfrak{S}_2$, i.e. within the scene surfaces and then $\chi_2 = 1$, see (23). So one can expand integration to the infinite space domain:

$$\begin{aligned}
&\int g(\boldsymbol{z}_2, L(\boldsymbol{z}_1, \boldsymbol{z}_2))\tilde{f}(\overrightarrow{\boldsymbol{z}_2\boldsymbol{z}_1}, \overrightarrow{\boldsymbol{z}_0\boldsymbol{z}_1}, \boldsymbol{z}_1)\left|(\boldsymbol{n}(\boldsymbol{z}_1)\cdot\overrightarrow{\boldsymbol{z}_2\boldsymbol{z}_1})\right| s(\boldsymbol{z}_1, \boldsymbol{z}_2)d\boldsymbol{z}_2 \\
&\approx a\int g(\boldsymbol{z}_2, L_{max})\tilde{f}(\overrightarrow{\boldsymbol{z}_2\boldsymbol{z}_1}, \overrightarrow{\boldsymbol{z}_0\boldsymbol{z}_1}, \boldsymbol{z}_1)\left|(\boldsymbol{n}(\boldsymbol{z}_1)\cdot\overrightarrow{\boldsymbol{z}_2\boldsymbol{z}_1})\right| s(\boldsymbol{z}_1, \boldsymbol{z}_2)d\boldsymbol{z}_2 \\
&+(1-a)\int g(\boldsymbol{z}_2, L_{min})\tilde{f}(\overrightarrow{\boldsymbol{z}_2\boldsymbol{z}_1}, \overrightarrow{\boldsymbol{z}_0\boldsymbol{z}_1}, \boldsymbol{z}_1)\left|(\boldsymbol{n}(\boldsymbol{z}_1)\cdot\overrightarrow{\boldsymbol{z}_2\boldsymbol{z}_1})\right| s(\boldsymbol{z}_1, \boldsymbol{z}_2)d\boldsymbol{z}_2
\end{aligned}$$

Here $s(\boldsymbol{z}_1, \boldsymbol{z}_2)d\boldsymbol{z}_2$ is nothing but the differential of the solid angle $d^2(\overrightarrow{\boldsymbol{z}_2\boldsymbol{z}_1})$, so if $g()$ does not depend on $\boldsymbol{z}_2$ explicitly, i.e. is of the form $g(\overrightarrow{\boldsymbol{z}_2\boldsymbol{z}_1}, L(\boldsymbol{z}_1, \boldsymbol{z}_2))$ we can replace integration over space with integration over the hemisphere of directions, so

$$\begin{aligned}
&\int g(\overrightarrow{\boldsymbol{z}_2\boldsymbol{z}_1}, L(\boldsymbol{z}_1, \boldsymbol{z}_2))\tilde{f}(\overrightarrow{\boldsymbol{z}_2\boldsymbol{z}_1}, \overrightarrow{\boldsymbol{z}_0\boldsymbol{z}_1}, \boldsymbol{z}_1)\left|(\boldsymbol{n}(\boldsymbol{z}_1)\cdot\overrightarrow{\boldsymbol{z}_2\boldsymbol{z}_1})\right| s(\boldsymbol{z}_1, \boldsymbol{z}_2)d\boldsymbol{z}_2 \\
&\approx a\int g(\boldsymbol{v}, L_{max})\tilde{f}(\boldsymbol{v}, \overrightarrow{\boldsymbol{z}_0\boldsymbol{z}_1}, \boldsymbol{z}_1)\left|(\boldsymbol{n}(\boldsymbol{z}_1)\cdot\boldsymbol{v})\right| d^2\boldsymbol{v} \quad (32) \\
&+(1-a)\int g(\boldsymbol{v}, L_{max})\tilde{f}(\boldsymbol{v}, \overrightarrow{\boldsymbol{z}_0\boldsymbol{z}_1}, \boldsymbol{z}_1)\left|(\boldsymbol{n}(\boldsymbol{z}_1)\cdot\boldsymbol{v})\right| d^2\boldsymbol{v}
\end{aligned}$$

A.1 Calculation of $\widetilde{\Gamma}_2$

We use eq. (22) for the assumed case $\chi_2 = 1$. Using approximation (32) and assuming $c(\boldsymbol{z}_1, \boldsymbol{z}_2)$ is constant one obtains from it

$$\begin{aligned}\widetilde{\Gamma}_2 &\approx \frac{L_{max} a I_2(c+n_F L_{max}) + L_{min}(1-a) I_2(c+n_F L_{min})}{a\frac{c+n_F L_{max} I_2(c+n_F L_{max})}{c+n_F L_{max}} + (1-a)\frac{c+n_F L_{min} I_2(c+n_F L_{min})}{c+n_F L_{min}}} \\ &= \frac{L_{max} a I_2(c+n_F L_{max}) + L_{min}(1-a) I_2(c+n_F L_{min})}{a\left(c\frac{1-I_2(c+n_F L_{max})}{c+n_F L_{max}} + I_2(c+n_F L_{max})\right) + (1-a)\left(c\frac{1-I_2(c+n_F L_{min})}{c+n_F L_{min}} + I_2(c+n_F L_{min})\right)} \\ &= \frac{L_{max} a I_2(c+n_F L_{max}) + L_{min}(1-a) I_2(c+n_F L_{min})}{a(cI_1(c+n_F L_{max}) + I_2(c+n_F L_{max})) + (1-a)(cI_1(c+n_F L_{min}) + I_2(c+n_F L_{min}))}\end{aligned} \tag{33}$$

where

$$I_m(\beta) \equiv \int \frac{\tilde{f}^m(\boldsymbol{v}, \boldsymbol{u}, \boldsymbol{z}_1)}{\tilde{f}(\boldsymbol{v}, \boldsymbol{u}, \boldsymbol{z}_1) + \beta} \left|(\boldsymbol{n}(\boldsymbol{z}_1) \cdot \boldsymbol{v})\right| d^2\boldsymbol{v} \tag{34}$$

and

$$I_1(\beta) = \frac{1 - I_2(\beta)}{\beta}$$

A.2 Calculation of $\mathcal{V}$

Substituting into (17) our high-frequency luminance and the optimal weights $\omega_{3,2} = \frac{\tilde{f}(\overrightarrow{\boldsymbol{z}_2\boldsymbol{z}_1}, \overrightarrow{\boldsymbol{z}_0\boldsymbol{z}_1}, \boldsymbol{z}_1) + n_F \widetilde{\Gamma}_2(\boldsymbol{z}_0, \boldsymbol{z}_1)}{\tilde{f}(\overrightarrow{\boldsymbol{z}_2\boldsymbol{z}_1}, \overrightarrow{\boldsymbol{z}_0\boldsymbol{z}_1}, \boldsymbol{z}_1) + c + n_F L(\boldsymbol{z}_1, \boldsymbol{z}_2)}$ and $\omega_{3,1} = 1 - \omega_{3,2}$ (where $\widetilde{\Gamma}_2$ and c are independent of $\boldsymbol{z}_2$), then expanding the squares and applying approximation (32) to the resulting integrals, one after some simple but tedious algebra comes to

$$\begin{aligned}\mathcal{V}(\boldsymbol{z}_0, \boldsymbol{z}_1) \approx\; & E^2(\boldsymbol{z}_0, \boldsymbol{z}_1)\left((L_{max} - \widetilde{\Gamma}_2)^2 + n_F^{-1} c L_{max}\right) a J_3(c + n_F L_{max}) \\ & + E^2(\boldsymbol{z}_0, \boldsymbol{z}_1)\left(c + n_F L_{max}\right) c \widetilde{\Gamma}_2^2 a J_1(c + n_F L_{max}) \\ & + E^2(\boldsymbol{z}_0, \boldsymbol{z}_1) a J_2(c + n_F L_{max}) \\ & \times \left(n_F^{-1} L_{max}\left(c + n_F L_{max} - n_F \widetilde{\Gamma}_2\right)^2 + 2c\widetilde{\Gamma}_2 L_{max} - 2c(L_{max} - \widetilde{\Gamma}_2)\widetilde{\Gamma}_2\right) \\ & + E^2(\boldsymbol{z}_0, \boldsymbol{z}_1)\left((L_{min} - \widetilde{\Gamma}_2)^2 + n_F^{-1} c L_{min}\right)(1-a) J_3(c + n_F L_{min}) \\ & + E^2(\boldsymbol{z}_0, \boldsymbol{z}_1)\left(c + n_F L_{min}\right) c \widetilde{\Gamma}_2^2 (1-a) J_1(c + n_F L_{min}) \\ & + E^2(\boldsymbol{z}_0, \boldsymbol{z}_1)(1-a) J_2(c + n_F L_{min}) \\ & \times \left(n_F^{-1} L_{min}\left(c + n_F L_{min} - n_F \widetilde{\Gamma}_2\right)^2 + 2c\widetilde{\Gamma}_2 L_{min} - 2c(L_{min} - \widetilde{\Gamma}_2)\widetilde{\Gamma}_2\right)\end{aligned}$$

where

$$J_m(\beta) \equiv \int \frac{\tilde{f}^m(\boldsymbol{v},\boldsymbol{u},\boldsymbol{z}_1)}{\left(\tilde{f}(\boldsymbol{v},\boldsymbol{u},\boldsymbol{z}_1)+\beta\right)^2} \left|(\boldsymbol{n}(\boldsymbol{z}_1)\cdot\boldsymbol{v})\right| d^2\boldsymbol{v} \tag{35}$$

The long expression for $\mathcal{V}$ includes two functionally identical groups, yet one for L_{min} and another for L_{max}. Therefore denoting the group

$$\begin{aligned} T(c,L,\widetilde{\Gamma}_2) &\equiv c\widetilde{\Gamma}_2^2(c+n_F L)J_1(c+n_F L) \\ &\quad + \left(n_F^{-1}L\left(c+n_F L-n_F\widetilde{\Gamma}_2\right)^2+2c\widetilde{\Gamma}_2^2\right)J_2(c+n_F L) \\ &\quad + \left(n_F^{-1}L\left(c+n_F L-n_F\widetilde{\Gamma}_2\right)+\widetilde{\Gamma}_2\left(\widetilde{\Gamma}_2-L\right)\right)J_3(c+n_F L) \\ &= \left(L-\widetilde{\Gamma}_2\right)^2\left(\left(c+n_F L\right)J_2+J_3\right) \\ &\quad + n_F^{-1}c\left(n_F\widetilde{\Gamma}_2^2\left(\left(c+n_F L\right)J_1+J_2\right)+L\left(\left(c+n_F L\right)J_2+J_3\right)\right) \end{aligned}$$

one can write

$$\mathcal{V}(\boldsymbol{z}_0,\boldsymbol{z}_1) = E^2(\boldsymbol{z}_0,\boldsymbol{z}_1)\left(aT(c,L_{max},\widetilde{\Gamma}_2)+(1-a)T(c,L_{min},\widetilde{\Gamma}_2)\right)$$

where $\widetilde{\Gamma}_2$ is given by (33).

Rewriting T in the form

$$\begin{aligned} T(c,L,\widetilde{\Gamma}_2) &= \left(L-\widetilde{\Gamma}_2\right)^2\left(\left(c+n_F L\right)J_2+J_3\right) \\ &\quad + n_F^{-1}c\left(n_F\widetilde{\Gamma}_2^2\left(\left(c+n_F L\right)J_1+J_2\right)+L\left(\left(c+n_F L\right)J_2+J_3\right)\right) \end{aligned}$$

and using the identities

$$\begin{aligned} \beta J_2(\beta)+J_3(\beta) &= I_2(\beta) \\ \beta J_1(\beta)+J_2(\beta) &= I_1(\beta) \end{aligned}$$

taking into account (34) we come to

$$
\begin{aligned}
T(c, L, \widetilde{\Gamma}_2) &= \left(L - \widetilde{\Gamma}_2\right)^2 I_2(c + n_F L) \\
&\quad + n_F^{-1} c \left(n_F \widetilde{\Gamma}_2^2 \frac{1 - I_2(c + n_F L)}{c + n_F L} + L I_2(c + n_F L) \right) \\
&= \left\{ \left(L - \widetilde{\Gamma}_2\right)^2 + c \left(n_F^{-1} L - \frac{\widetilde{\Gamma}_2^2}{c + n_F L} \right) \right\} I_2(c + n_F L) + c \frac{\widetilde{\Gamma}_2^2}{c + n_F L} \\
&= \left\{ \left(L - \widetilde{\Gamma}_2\right)^2 + c n_F^{-1} L \right\} I_2(c + n_F L) + c \widetilde{\Gamma}_2^2 \frac{1 - I_2(c + n_F L)}{c + n_F L} \\
&= n_F^{-1} L (c + n_F L) I_2(c + n_F L) - 2 \widetilde{\Gamma}_2 L I_2(c + n_F L) \\
&\quad + \widetilde{\Gamma}_2^2 \frac{c + n_F L I_2(c + n_F L)}{c + n_F L} \\
&= \frac{L n_F^{-1} \left(c + n_F L - n_F \widetilde{\Gamma}_2 \right)^2 I_2(c + n_F L) + c \widetilde{\Gamma}_2^2}{c + n_F L}
\end{aligned}
$$

For small $c \to 0$

$$
\widetilde{\Gamma}_2 \approx \frac{L_{max} a I_2(n_F L_{max}) + L_{min}(1 - a) I_2(n_F L_{min})}{a I_2(n_F L_{max}) + (1 - a) I_2(n_F L_{min})}
$$

$$
T(c \to 0, L, \widetilde{\Gamma}_2) \approx \left(L - \widetilde{\Gamma}_2\right)^2 I_2(n_F L)
$$

$$
\begin{aligned}
\mathcal{V}(\boldsymbol{z}_0, \boldsymbol{z}_1) &\approx E^2(\boldsymbol{z}_0, \boldsymbol{z}_1) \\
&\quad \times \left(a \left(L_{max} - \widetilde{\Gamma}_2\right)^2 I_2(n_F L_{max}) + (1 - a) \left(L_{min} - \widetilde{\Gamma}_2\right)^2 I_2(n_F L_{min}) \right) \\
L_{max} - \widetilde{\Gamma}_2 &\approx (1 - a)(L_{max} - L_{min}) \frac{I_2(n_F L_{min})}{a I_2(n_F L_{max}) + (1 - a) I_2(n_F L_{min})} \\
L_{min} - \widetilde{\Gamma}_2 &\approx a (L_{max} - L_{min}) \frac{I_2(n_F L_{max})}{a I_2(n_F L_{max}) + (1 - a) I_2(n_F L_{min})}
\end{aligned}
$$

thus now

$$
\mathcal{V}(\boldsymbol{z}_0, \boldsymbol{z}_1) \approx E^2(\boldsymbol{z}_0, \boldsymbol{z}_1) a(1 - a) \frac{I_2(n_F L_{max}) I_2(n_F L_{min})}{a I_2(n_F L_{max}) + (1 - a) I_2(n_F L_{min})} (L_{max} - L_{min})^2
$$

Further calculations are done for the case of normal incidence and Gaussian BDF

$$\tilde{f}(\boldsymbol{v},\boldsymbol{u},\boldsymbol{z}_1) = \frac{1}{\pi w^2} e^{-\left(\frac{\vartheta}{w}\right)^2}$$

where ϑ is the angle between the outgoing ray and mirror reflection of the incident ray, and w is BDF width.

Now one has

$$\begin{aligned} I_2(\beta) &= 2\pi \int_0^{\pi/2} \frac{\left(\frac{1}{\pi w^2} e^{-\left(\frac{\vartheta}{w}\right)^2}\right)^2}{\beta + \frac{1}{\pi w^2} e^{-\left(\frac{\vartheta}{w}\right)^2}} \vartheta d\vartheta \\ &\approx 2\int_0^{\infty} \frac{\left(e^{-t^2}\right)^2}{\pi w^2 \beta + e^{-t^2}} t dt \\ &= 1 + (\pi w^2 \beta) \ln \frac{(\pi w^2 \beta)}{1 + (\pi w^2 \beta)} \\ &\approx \frac{1}{1 + 2.05 \pi w^2 \beta} \end{aligned}$$

and

$$\mathcal{V}(\boldsymbol{z}_0,\boldsymbol{z}_1) \approx E^2(\boldsymbol{z}_0,\boldsymbol{z}_1) a(1-a)(L_{max} - L_{min})^2 \times \frac{\left(1 + (\pi w^2 n_F L_{max}) \ln \frac{(\pi w^2 n_F L_{max})}{1+(\pi w^2 n_F L_{max})}\right)\left(1 + (\pi w^2 n_F L_{min}) \ln \frac{(\pi w^2 n_F L_{min})}{1+(\pi w^2 n_F L_{min})}\right)}{1 + \pi w^2 \left\{a(n_F L_{max}) \ln \frac{(\pi w^2 n_F L_{max})}{1+(\pi w^2 n_F L_{max})} + (1-a)(n_F L_{min}) \ln \frac{(\pi w^2 n_F L_{min})}{1+(\pi w^2 n_F L_{min})}\right\}}$$

For sharp BDF, when $w^2 n_F L_{max} \to 0$, $I_2(\beta) \to 1$:

$$\begin{aligned} \widetilde{\Gamma}_2 &\approx L_{max} a + L_{min}(1-a) \\ \omega_{3,2} &\approx \frac{\tilde{f}(\overrightarrow{\boldsymbol{z}_2\boldsymbol{z}_1}, \overrightarrow{\boldsymbol{z}_0\boldsymbol{z}_1}, \boldsymbol{z}_1) + n_F\left(L_{max} a + L_{min}(1-a)\right)}{\tilde{f}(\overrightarrow{\boldsymbol{z}_2\boldsymbol{z}_1}, \overrightarrow{\boldsymbol{z}_0\boldsymbol{z}_1}, \boldsymbol{z}_1) + c + n_F L(\boldsymbol{z}_1,\boldsymbol{z}_2)} \approx 1 \\ \mathcal{V}(\boldsymbol{z}_0,\boldsymbol{z}_1) &\approx E^2(\boldsymbol{z}_0,\boldsymbol{z}_1) a(1-a)(L_{max} - L_{min})^2 \\ \frac{n_F \mathcal{V}(\boldsymbol{z}_0,\boldsymbol{z}_1)}{E^2(\boldsymbol{z}_0) L(\boldsymbol{z}_0,\boldsymbol{z}_1)} &\approx \frac{E(\boldsymbol{z}_0,\boldsymbol{z}_1)}{E(\boldsymbol{z}_0)} n_F \frac{a(1-a)(L_{max} - L_{min})^2}{a L_{max} + (1-a) L_{min}} \approx \frac{E(\boldsymbol{z}_0,\boldsymbol{z}_1)}{E(\boldsymbol{z}_0)}(1-a) n_F L_{max} \end{aligned}$$

It should be noted that the maximum of $\frac{n_F \mathcal{V}(\boldsymbol{z}_0,\boldsymbol{z}_1)}{E^2(\boldsymbol{z}_0) L(\boldsymbol{z}_0,\boldsymbol{z}_1)}$ is for $a = \frac{\sqrt{L_{min}}}{\sqrt{L_{max}} + \sqrt{L_{min}}}$ and equals

$$\frac{n_F \mathcal{V}(\boldsymbol{z}_0, \boldsymbol{z}_1)}{E^2(\boldsymbol{z}_0) L(\boldsymbol{z}_0, \boldsymbol{z}_1)} \approx \frac{E(\boldsymbol{z}_0, \boldsymbol{z}_1)}{E(\boldsymbol{z}_0)} n_F \frac{(L_{max} - L_{min})^2}{\left(\sqrt{L_{max}} + \sqrt{L_{min}}\right)^2}$$

For high contrast, i.e. $L_{max} \gg L_{min}$, this becomes

$$\frac{n_F \mathcal{V}(\boldsymbol{z}_0, \boldsymbol{z}_1)}{E^2(\boldsymbol{z}_0) L(\boldsymbol{z}_0, \boldsymbol{z}_1)} \approx \frac{E(\boldsymbol{z}_0, \boldsymbol{z}_1)}{E(\boldsymbol{z}_0)} n_F L_{max}$$

In the case of not sharp BDF, when $w n_F L$ is large, $I_2(\beta) \to \frac{1}{2\pi w^2 \beta}$:

$$\begin{aligned}
\widetilde{\Gamma}_2 &\approx \frac{L_{max} L_{min}}{a L_{min} + (1-a) L_{max}} \approx \frac{L_{min}}{1-a} \\
\omega_{3,2} &\approx \frac{\tilde{f}(\overrightarrow{\boldsymbol{z}_2 \boldsymbol{z}_1}, \overrightarrow{\boldsymbol{z}_0 \boldsymbol{z}_1}, \boldsymbol{z}_1) + \frac{n_F L_{min}}{1-a}}{\tilde{f}(\overrightarrow{\boldsymbol{z}_2 \boldsymbol{z}_1}, \overrightarrow{\boldsymbol{z}_0 \boldsymbol{z}_1}, \boldsymbol{z}_1) + c + n_F L(\boldsymbol{z}_1, \boldsymbol{z}_2)} \approx 0 \\
\mathcal{V}(\boldsymbol{z}_0, \boldsymbol{z}_1) &\approx E^2(\boldsymbol{z}_0, \boldsymbol{z}_1) a(1-a) \frac{1}{2} \frac{1}{\pi w^2} \frac{n_F^{-1} (L_{max} - L_{min})^2}{L_{max} - a(L_{max} - L_{min})} \\
&\approx E^2(\boldsymbol{z}_0, \boldsymbol{z}_1) a(1-a) \frac{1}{2\pi w^2} n_F^{-1} L_{max} \\
\frac{n_F \mathcal{V}(\boldsymbol{z}_0, \boldsymbol{z}_1)}{E^2(\boldsymbol{z}_0) L(\boldsymbol{z}_0, \boldsymbol{z}_1)} &\approx \frac{E(\boldsymbol{z}_0, \boldsymbol{z}_1)}{E(\boldsymbol{z}_0)} (1-a) \frac{1}{2\pi w^2}
\end{aligned}$$

A.3 The case of uniform luminance

Now let's $L_{max} = L_{min} = L$.

$$\widetilde{\Gamma}_2 = (c + n_F L) \frac{L I_2(c + n_F L)}{c + n_F L I_2(c + n_F L)}$$

$$\begin{aligned}
T(c, L, \widetilde{\Gamma}_2) &= \frac{L n_F^{-1} \left(c + n_F L - n_F \widetilde{\Gamma}_2\right)^2 I_2(c + n_F L) + c \widetilde{\Gamma}_2^2}{c + n_F L} \\
&= c n_F^{-2} (c + n_F L) \frac{n_F L I_2(c + n_F L)}{c + n_F L I_2(c + n_F L)}
\end{aligned}$$

$$\mathcal{V}(\boldsymbol{z}_0, \boldsymbol{z}_1) = E^2(\boldsymbol{z}_0, \boldsymbol{z}_1) c n_F^{-2} (c + n_F L) \frac{n_F L I_2(c + n_F L)}{c + n_F L I_2(c + n_F L)}$$

For sharp BDF, when $w n_F L \to 0$, $I_2(\beta) \to 1$:

$$\begin{aligned}
\omega_{3,2} &= 1 \\
\mathcal{V}(\boldsymbol{z}_0, \boldsymbol{z}_1) &= E^2(\boldsymbol{z}_0, \boldsymbol{z}_1) c n_F^{-1} L \\
\frac{n_F \mathcal{V}(\boldsymbol{z}_0, \boldsymbol{z}_1)}{E^2(\boldsymbol{z}_0) L(\boldsymbol{z}_0, \boldsymbol{z}_1)} &\approx \frac{n_F E(\boldsymbol{z}_0, \boldsymbol{z}_1) c n_F^{-1} L}{E(\boldsymbol{z}_0) L} = \frac{E(\boldsymbol{z}_0, \boldsymbol{z}_1)}{E(\boldsymbol{z}_0)} c
\end{aligned}$$

In the case of not sharp BDF, when $w n_F L$ is large, $I_2(\beta) \to \frac{1}{2\pi w^2 \beta}$:

$$\begin{aligned}
\widetilde{\Gamma}_2 &= \frac{L}{2\pi w^2 c + \frac{n_F L}{c + n_F L}} \approx \frac{L}{1 + 2\pi w^2 c} \\
\omega_{3,2} &= \frac{\tilde{f}(\overrightarrow{\boldsymbol{z}_2 \boldsymbol{z}_1}, \overrightarrow{\boldsymbol{z}_0 \boldsymbol{z}_1}, \boldsymbol{z}_1) + \frac{n_F L}{2\pi w^2 c + \frac{n_F L}{c + n_F L}}}{\tilde{f}(\overrightarrow{\boldsymbol{z}_2 \boldsymbol{z}_1}, \overrightarrow{\boldsymbol{z}_0 \boldsymbol{z}_1}, \boldsymbol{z}_1) + c + n_F L} \approx \frac{1}{1 + 2\pi w^2 c} \\
\mathcal{V}(\boldsymbol{z}_0, \boldsymbol{z}_1) &\approx E^2(\boldsymbol{z}_0, \boldsymbol{z}_1) c n_F^{-2} \frac{n_F L}{2\pi w^2 c + \frac{n_F L}{c + n_F L}} \approx E^2(\boldsymbol{z}_0, \boldsymbol{z}_1) c n_F^{-2} \frac{n_F L}{2\pi w^2 c + 1} \\
&\xrightarrow{c \to \infty} E^2(\boldsymbol{z}_0, \boldsymbol{z}_1) \frac{1}{2\pi w^2} n_F^{-1} L \\
\frac{n_F \mathcal{V}(\boldsymbol{z}_0, \boldsymbol{z}_1)}{E^2(\boldsymbol{z}_0) L(\boldsymbol{z}_0, \boldsymbol{z}_1)} &\approx \frac{E(\boldsymbol{z}_0, \boldsymbol{z}_1)}{E(\boldsymbol{z}_0)} \frac{c}{2\pi w^2 c + 1} \xrightarrow{c \to \infty} \frac{E(\boldsymbol{z}_0, \boldsymbol{z}_1)}{E(\boldsymbol{z}_0)} \frac{1}{2\pi w^2}
\end{aligned}$$

References

[1] V. Frolov, A. Voloboy, S. Ershov, and V. Galaktionov, "Light transport in realistic rendering: state-of-the-art simulation methods," *Programming and Computer Software*, vol. 47, no. 4, pp. 298–326, 2021.

[2] M. Šik and J. Křivánek, "Survey of Markov chain Monte Carlo methods in light transport simulation," *IEEE Trans. Visual. Comput. Graph.*, vol. 26, no. 4, pp. 1821–1840, 2018.

[3] M. Pharr and G. Humphreys, *Physically Based Rendering, Second Edition: From Theory To Implementation*, 2nd ed. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 2010.

[4] H. W. Jensen and P. Christensen, "High quality rendering using ray tracing and photon mapping," in *ACM SIGGRAPH 2007 Courses*. New York, NY, USA: ACM, 2007. [Online]. Available: http://doi.acm.org/10.1145/1281500.1281593

[5] J. Vorba, "Bidirectional photon mapping," in *Proceedings of CESCG 2011: The 15th Central European Seminar on Computer Graphics*. Prague: Charles University, 2011, pp. 25–32. [Online]. Available: https://cgg.mff.cuni.cz/ jaroslav/papers/2011-bdpm/vorba2011-bdpm.pdf

[6] S. V. Ershov, D. D. Zhdanov, and A. G. Voloboy, "Estimation of noise in calculation of scattering medium luminance by MCRT," *Mathematica Montisnigri*, vol. 45, pp. 60–73, 2019. [Online]. Available: http://doi.acm.org/10.20948/mathmontis-2019-45-5

[7] I. Georgiev, J. Křivánek, T. Davidovič, and P. Slusallek, "Light transport simulation with vertex connection and merging," *ACM Trans. Graph.*, vol. 31, no. 6, pp. 192:1–192:10, 2012. [Online]. Available: http://doi.acm.org/10.1145/2366145.2366211

[8] T. Hachisuka, J. Pantaleoni, and H. W. Jensen, "A path space extension for robust light transport simulation," *ACM Trans. Graph.*, vol. 31, no. 6, pp. 191:1–191:10, 2012.

[9] M. Sbert, V. Havran, and L. Szirmay-Kalos, "Multiple importance sampling revisited: breaking the bounds," *EURASIP Journal on Advances in Signal Processing*, vol. 2018, pp. 1–15, 2018.

[10] E. Veach, "Robust Monte-Carlo methods for light transport simulation," 1997. [Online]. Available: http://graphics.stanford.edu/papers/veach_thesis/thesis.pdf

[11] I. Kondapaneni, P. Vévoda, P. Grittmann, T. Skřivan, P. Slusallek, and J. Křivánek, "Optimal multiple importance sampling," *ACM Transactions on Graphics (Proceedings of SIGGRAPH 2019)*, vol. 38, no. 4, pp. 37:1–37:14, 2019.

[12] A. Keller, P. Grittmann, J. Vorba, I. Georgiev, M. Šik, E. d'Eon, P. Gautron, P. Vévoda, and I. Kondapaneni, "Advances in Monte Carlo rendering: The legacy of Jaroslav Křivánek," in *ACM SIGGRAPH 2020 Courses*, New York, NY, USA, 2020, pp. 1–366. [Online]. Available: https://doi.org/10.1145/3388769.3407458

[13] S. Ershov and A. Voloboy, "Calculation of MIS weights for bidirectional path tracing with photon maps in presence of direct illumination," *Mathematica Montisnigri*, vol. 48, pp. 86–102, 2020. [Online]. Available: http://doi.acm.org/10.20948/mathmontis-2020-48-8

[14] S. V. Ershov, E. D. Birukov, and A. G. Voloboy, "Efficient calculation of the optimal MIS weights in bi-directional ray tracing with photon maps," *Keldysh Institute Preprints*, no. 107, 2020. [Online]. Available: https://doi.org/10.20948/prepr-2020-107

[15] S. Ershov, M. Kopylov, and A. Voloboy, "Optimal weights for bidirectional ray tracing with photon maps while mixing 3 strategies," *Keldysh Institute Preprints*, no. 89, 2021. [Online]. Available: https://doi.org/10.20948/prepr-2021-89-e

[16] E. A. Kopylov, A. Khodulev, and V. Volevich, "The comparison of illumination maps technique in computer graphics software," 1998. [Online]. Available: https://api.semanticscholar.org/CorpusID:14690221